Distributed convergence to Nash equilibria in two-network zero-sum games ⋆

B. Gharesifard [a] J. Cortés [b]

[a] *Coordinated Science Laboratory, University of Illinois at Urbana-Champaign, IL, 61801, USA*

[b] *Department of Mechanical and Aerospace Engineering, University of California, San Diego, CA, 92093, USA*

Abstract

This paper considers a class of strategic scenarios in which two networks of agents have opposing objectives with regards to the optimization of a common objective function. In the resulting zero-sum game, individual agents collaborate with neighbors in their respective network and have only partial knowledge of the state of the agents in the other network. For the case when the interaction topology of each network is undirected, we synthesize a distributed saddle-point strategy and establish its convergence to the Nash equilibrium for the class of strictly concave-convex and locally Lipschitz objective functions. We also show that this dynamics does not converge in general if the topologies are directed. This justifies the introduction, in the directed case, of a generalization of this distributed dynamics which we show converges to the Nash equilibrium for the class of strictly concave-convex differentiable functions with globally Lipschitz gradients. The technical approach combines tools from algebraic graph theory, nonsmooth analysis, set-valued dynamical systems, and game theory.

Key words: adversarial networks, distributed algorithms, zero-sum game, saddle-point dynamics, Nash equilibria

1 Introduction

Recent years have seen an increasing interest on networked strategic scenarios where agents may cooperate or compete with each other towards the achievement of some objective, interact across different layers, have access to limited information, and are subject to evolving interaction topologies. This paper is a contribution to this body of work. Specifically, we consider a class of strategic scenarios in which two networks of agents are involved in a zero-sum game. We assume that the objective function can be decomposed as a sum of concave-convex functions and that the networks have opposing objectives regarding its optimization. Agents collaborate with the neighbors in their own network and have partial information about the state of the agents in the other network. Such scenarios are challenging because information is spread across the agents and possibly multiple layers, and networks, by themselves, are not the decision makers. Our aim is to design a distributed coordination algorithm that can be used by the agents to converge to the Nash equilibrium. Note that, for a 2-player zero-sum game of the type considered here, a pure Nash equilibrium corresponds to a saddle point of the objective function.

Literature review. Multiple scenarios involving networked systems and intelligent adversaries in sensor networks, filtering, finance, and wireless communications (Kim and Boyd, 2008; Wan and Lemmon, 2009) can be cast into the strategic framework described above. In such scenarios, the network objective arises as a result of the aggregation of agent-to-agent adversarial interactions regarding a common goal, and information is naturally distributed among the agents. The present work has connections with the literature on distributed optimization and zero-sum games. The distributed optimization of a sum of convex functions has been intensively studied in recent years, see e.g. (Nedic and Ozdaglar, 2009; Wan and Lemmon, 2009; Johansson et al., 2009; Zhu and Martínez, 2012). These works build on consensus-based dynamics (Olfati-Saber et al., 2007; Ren and Beard, 2008; Bullo et al., 2009; Mesbahi and Egerstedt, 2010) to find the solutions of the optimization problem in a variety of scenarios and are designed in discrete

⋆ Incomplete versions of this paper were presented as (Gharesifard and Cortés, 2012a) in the American Control Conference 2012 and as (Gharesifard and Cortés, 2012b) in the IEEE Control and Decision Conference 2012. This work was performed while B. Gharesifard was a postdoctoral researcher at the University of California, San Diego.

Email addresses: bgharesi@illinois.edu (B. Gharesifard), cortes@ucsd.edu (J. Cortés).

time. Exceptions include (Wang and Elia, 2010, 2011) on continuous-time distributed optimization on undirected networks and (Gharesifard and Cortés, 2012c) on directed networks.

Regarding zero-sum games, the works (Arrow et al., 1958; Maistroskii, 1977; Nedic and Ozdgalar, 2009) study the convergence of discrete-time subgradient dynamics to a saddle point. Continuous-time best-response dynamics for zero-sum games converges to the set of Nash equilibria for both convex-concave (Hofbauer and Sorin, 2006) and quasiconvex-quasiconcave (Barron et al., 2010) functions. Under strict convexity-concavity assumptions, continuous-time subgradient flow dynamics converges to a saddle point (Arrow et al., 1951, 1958). Asymptotic convergence is also guaranteed when the Hessian of the objective function is positive definite in one argument and the function is linear in the other (Arrow et al., 1958; Feijer and Paganini, 2010). The distributed computation of Nash equilibria in noncooperative games, where all players are adversarial, has been investigated under different assumptions. The algorithm in (Li and Başar, 1987) relies on all-to-all communication and does not require players to know each other's payoff functions (which must be strongly convex). In (Frihauf et al., 2012; Stankovic et al., 2012), players are unaware of their own payoff functions but have access to the payoff value of an action once it has been executed. These works design distributed strategies based on extremum seeking techniques to seek the set of Nash equilibria.

Statement of contributions. We introduce the problem of distributed convergence to Nash equilibria for two networks engaged in a strategic scenario. The networks aim to either maximize or minimize a common objective function which can be written as a sum of concave-convex functions. Individual agents collaborate with neighbors in their respective network and have partial knowledge of the state of the agents in the other one. Our first contribution is the introduction of an aggregate objective function for each network which depends on the interaction topology through its Laplacian and the characterization of a family of points with a saddle property for the pair of functions. We show the correspondence between these points and the Nash equilibria of the overall game. When the graphs describing the interaction topologies within each network are undirected, the gradients of these aggregate objective functions are distributed. Building on this observation, our second contribution is the synthesis of a consensus-based saddle-point strategy for adversarial networks with undirected topologies. We show that the proposed dynamics is guaranteed to asymptotically converge to the Nash equilibrium for the class of strictly concave-convex and locally Lipschitz objective functions. Our third contribution focuses on the directed case. We show that the transcription of the saddle-point dynamics to directed topologies fails to converge in general. This leads us to propose a generalization of the dynamics, for strongly connected weight-balanced topologies, that incorporates a design parameter. We show that, by appropriately choosing this parameter, the new dynamics asymptotically converges to the Nash equilibrium for the class of strictly concave-convex differentiable objective functions with globally Lipschitz gradients. The technical approach employs notions and results from algebraic graph theory, nonsmooth and convex analysis, set-valued dynamical systems, and game theory. As an intermediate result in our proof strategy for the directed case, we provide a generalization of the known characterization of cocoercivity of concave functions to concave-convex functions.

The results of this paper can be understood as a generalization to competing networks of the results we obtained in (Gharesifard and Cortés, 2012c) for distributed optimization. This generalization is nontrivial because the payoff functions associated to individual agents now also depend on information obtained from the opposing network. This feature gives rise to a hierarchy of saddle-point dynamics whose analysis is technically challenging and requires, among other things, a reformulation of the problem as a constrained zero-sum game, a careful understanding of the coupling between the dynamics of both networks, and the generalization of the notion of cocoercivity to concave-convex functions.

Organization. Section 2 contains preliminaries on nonsmooth analysis, set-valued dynamical systems, graph theory, and game theory. In Section 3, we introduce the zero-sum game for two adversarial networks involved in a strategic scenario and introduce two novel aggregate objective functions. Section 4 presents our algorithm design and analysis for distributed convergence to Nash equilibrium when the network topologies are undirected. Section 5 presents our treatment for the directed case. Section 6 gathers our conclusions and ideas for future work. Appendix A contains the generalization to concave-convex functions of the characterization of cocoercivity of concave functions.

2 Preliminaries

We start with some notational conventions. Let $\mathbb{R}$, $\mathbb{R}_{\geq 0}$, $\mathbb{Z}$, $\mathbb{Z}_{\geq 1}$ denote the set of real, nonnegative real, integer, and positive integer numbers, respectively. We denote by $||\cdot||$ the Euclidean norm on $\mathbb{R}^d$, $d \in \mathbb{Z}_{\geq 1}$ and also use the short-hand notation $\mathbf{1}_d = (1, \dots, 1)^T$ and $\mathbf{0}_d = (0, \dots, 0)^T \in \mathbb{R}^d$. We let I_d denote the identity matrix in $\mathbb{R}^{d \times d}$. For matrices $A \in \mathbb{R}^{d_1 \times d_2}$ and $B \in \mathbb{R}^{e_1 \times e_2}$, $d_1, d_2, e_1, e_2 \in \mathbb{Z}_{\geq 1}$, we let $A \otimes B$ denote their Kronecker product. The function $f : \mathsf{X}_1 \times \mathsf{X}_2 \to \mathbb{R}$, with $\mathsf{X}_1 \subset \mathbb{R}^{d_1}$, $\mathsf{X}_2 \subset \mathbb{R}^{d_2}$ closed and convex, is *concave-convex* if it is concave in its first argument and convex in the second one (Rockafellar, 1997). A point $(x_1^*, x_2^*) \in \mathsf{X}_1 \times \mathsf{X}_2$ is a *saddle point* of f if $f(x_1, x_2^*) \leq f(x_1^*, x_2^*) \leq f(x_1^*, x_2)$ for all $x_1 \in \mathsf{X}_1$ and $x_2 \in \mathsf{X}_2$. Finally, a set-valued map $f : \mathbb{R}^d \rightrightarrows \mathbb{R}^d$ takes elements of $\mathbb{R}^d$ to subsets of $\mathbb{R}^d$.

2.1 Nonsmooth analysis

We recall some notions from nonsmooth analysis (Clarke, 1983). A function $f : \mathbb{R}^d \to \mathbb{R}$ is *locally Lipschitz* at $x \in \mathbb{R}^d$ if there exists a neighborhood $\mathcal{U}$ of x and $C_x \in \mathbb{R}_{\geq 0}$ such that $|f(y) - f(z)| \leq C_x||y - z||$, for $y, z \in \mathcal{U}$. f is locally Lipschitz on $\mathbb{R}^d$ if it is locally Lipschitz at x for all $x \in \mathbb{R}^d$ and *globally Lipschitz* on $\mathbb{R}^d$ if for all $y, z \in \mathbb{R}^d$ there exists $C \in \mathbb{R}_{\geq 0}$ such that $|f(y) - f(z)| \leq C||y - z||$. Locally Lipschitz functions are differentiable almost everywhere. The *generalized gradient* of f is

$$\partial f(x) = \mathrm{co}\Big\{ \lim_{k\to\infty} \nabla f(x_k) \mid x_k \to x, x_k \notin \Omega_f \cup S \Big\},$$

where Ω_f is the set of points where f fails to be differentiable and S is any set of measure zero.

Lemma 2.1 (Continuity of the generalized gradient map): *Let $f : \mathbb{R}^d \to \mathbb{R}$ be a locally Lipschitz function at $x \in \mathbb{R}^d$. Then the set-valued map $\partial f : \mathbb{R}^d \rightrightarrows \mathbb{R}^d$ is upper semicontinuous and locally bounded at $x \in \mathbb{R}^d$ and moreover, $\partial f(x)$ is nonempty, compact, and convex.*

For $f : \mathbb{R}^d \times \mathbb{R}^d \to \mathbb{R}$ and $z \in \mathbb{R}^d$, we let $\partial_x f(x, z)$ denote the generalized gradient of $x \mapsto f(x, z)$. Similarly, for $x \in \mathbb{R}^d$, we let $\partial_z f(x, z)$ denote the generalized gradient of $z \mapsto f(x, z)$. A point $x \in \mathbb{R}^d$ with $\mathbf{0} \in \partial f(x)$ is a *critical point* of f. A function $f : \mathbb{R}^d \to \mathbb{R}$ is *regular* at $x \in \mathbb{R}$ if for all $v \in \mathbb{R}^d$ the right directional derivative of f, in the direction of v, exists at x and coincides with the generalized directional derivative of f at x in the direction of v. We refer the reader to (Clarke, 1983) for definitions of these notions. A convex and locally Lipschitz function at x is regular (Clarke, 1983, Proposition 2.3.6). The notion of regularity plays an important role when considering sums of Lipschitz functions.

Lemma 2.2 (Finite sum of locally Lipschitz functions): *Let $\{f^i\}_{i=1}^n$ be locally Lipschitz at $x \in \mathbb{R}^d$. Then $\partial(\sum_{i=1}^n f^i)(x) \subseteq \sum_{i=1}^n \partial f^i(x)$, and equality holds if f^i is regular for $i \in \{1, \dots, n\}$.*

A locally Lipschitz and convex function f satisfies, for all $x, x' \in \mathbb{R}^d$ and $\xi \in \partial f(x)$, the *first-order condition* of convexity,

$$f(x') - f(x) \geq \xi \cdot (x' - x). \qquad (1)$$

2.2 Set-valued dynamical systems

Here, we recall some background on set-valued dynamical systems following Cortés (2008). A continuous-time set-valued dynamical system on $\mathsf{X} \subset \mathbb{R}^d$ is a differential inclusion

$$\dot{x}(t) \in \Psi(x(t)) \qquad (2)$$

where $t \in \mathbb{R}_{\geq 0}$ and $\Psi : \mathsf{X} \subset \mathbb{R}^d \rightrightarrows \mathbb{R}^d$ is a set-valued map. A solution to this dynamical system is an absolutely continuous curve $x : [0, T] \to \mathsf{X}$ which satisfies (2) almost everywhere. The set of equilibria of (2) is denoted by $\mathrm{Eq}(\Psi) = \{x \in \mathsf{X} \mid 0 \in \Psi(x)\}$.

Lemma 2.3 (Existence of solutions): *For $\Psi : \mathbb{R}^d \rightrightarrows \mathbb{R}^d$ upper semicontinuous with nonempty, compact, and convex values, there exists a solution to (2) from any initial condition.*

The LaSalle Invariance Principle for set-valued continuous-time systems is helpful to establish the asymptotic stability properties of systems of the form (2). A set $W \subset \mathsf{X}$ is *weakly positively invariant* with respect to Ψ if for any $x \in W$, there exists $\tilde{x} \in \mathsf{X}$ such that $\tilde{x} \in \Psi(x)$. The set W is *strongly positively invariant* with respect to Ψ if $\Psi(x) \subset W$, for all $x \in W$. Finally, the *set-valued Lie derivative* of a differentiable function $V : \mathbb{R}^d \to \mathbb{R}$ with respect to Ψ at $x \in \mathbb{R}^d$ is defined by $\widetilde{\mathcal{L}}_\Psi V(x) = \{v \cdot \nabla V(x) \mid v \in \Psi(x)\}$.

Theorem 2.4 (Set-valued LaSalle Invariance Principle): *Let $W \subset \mathsf{X}$ be a strongly positively invariant under (2) and $V : \mathsf{X} \to \mathbb{R}$ a continuously differentiable function. Suppose the evolutions of (2) are bounded and $\max \widetilde{\mathcal{L}}_\Psi V(x) \leq 0$ or $\widetilde{\mathcal{L}}_\Psi V(x) = \emptyset$, for all $x \in W$. If $S_{\Psi,V} = \{x \in \mathsf{X} \mid 0 \in \widetilde{\mathcal{L}}_\Psi V(x)\}$, then any solution $x(t)$, $t \in \mathbb{R}_{\geq 0}$, starting in W converges to the largest weakly positively invariant set M contained in $\bar{S}_{\Psi,V} \cap W$. When M is a finite collection of points, then the limit of each solution equals one of them.*

2.3 Graph theory

We present some basic notions from algebraic graph theory following the exposition in (Bullo et al., 2009). A *directed graph*, or simply *digraph*, is a pair $\mathcal{G} = (\mathcal{V}, \mathcal{E})$, where $\mathcal{V}$ is a finite set called the vertex set and $\mathcal{E} \subseteq \mathcal{V} \times \mathcal{V}$ is the edge set. A digraph is *undirected* if $(v, u) \in \mathcal{E}$ anytime $(u, v) \in \mathcal{E}$. We refer to an undirected digraph as a *graph*. A path is an ordered sequence of vertices such that any ordered pair of vertices appearing consecutively is an edge of the digraph. A digraph is *strongly connected* if there is a path between any pair of distinct vertices. For a graph, we refer to this notion simply as *connected*. A *weighted digraph* is a triplet $\mathcal{G} = (\mathcal{V}, \mathcal{E}, \mathsf{A})$, where $(\mathcal{V}, \mathcal{E})$ is a digraph and $\mathsf{A} \in \mathbb{R}_{\geq 0}^{n\times n}$ is the *adjacency matrix* of $\mathcal{G}$, with the property that $a_{ij} > 0$ if $(v_i, v_j) \in \mathcal{E}$ and $a_{ij} = 0$, otherwise. The weighted out-degree and in-degree of v_i, $i \in \{1, \dots, n\}$, are respectively, $d^{\mathrm{w}}_{\mathrm{out}}(v_i) = \sum_{j=1}^n a_{ij}$ and $d^{\mathrm{w}}_{\mathrm{in}}(v_i) = \sum_{j=1}^n a_{ji}$. The *weighted out-degree matrix* $\mathsf{D}_{\mathrm{out}}$ is the diagonal matrix defined by $(\mathsf{D}_{\mathrm{out}})_{ii} = d^{\mathrm{w}}_{\mathrm{out}}(i)$, for all $i \in \{1, \dots, n\}$. The *Laplacian* matrix is $\mathsf{L} = \mathsf{D}_{\mathrm{out}} - \mathsf{A}$. Note that $\mathsf{L}\mathbf{1}_n = 0$. If $\mathcal{G}$ is strongly connected, then zero is a simple eigenvalue of L. $\mathcal{G}$ is undirected if $\mathsf{L} = \mathsf{L}^T$ and *weight-balanced* if $d^{\mathrm{w}}_{\mathrm{out}}(v) = d^{\mathrm{w}}_{\mathrm{in}}(v)$, for all $v \in \mathcal{V}$. Equivalently, $\mathcal{G}$ is weight-balanced if and only if $\mathbf{1}_n^T\mathsf{L} = 0$ if and only if $\mathsf{L} + \mathsf{L}^T$ is positive semidefinite. Furthermore, if $\mathcal{G}$ is weight-balanced and strongly connected, then zero is a simple eigenvalue of $\mathsf{L} + \mathsf{L}^T$. Note that any undirected graph is weight-balanced.

2.4 Zero-sum games

We recall basic game-theoretic notions following Başar and Olsder (1999). An n-player game is a triplet $\mathbf{G} = (P, \mathsf{X}, U)$,

where P is the set of players with $|P| = n \in \mathbb{Z}_{\geq 2}$, $\mathsf{X} = \mathsf{X}_1 \times \ldots \times \mathsf{X}_n$, $\mathsf{X}_i \subset \mathbb{R}^{d_i}$ is the set of (pure) strategies of player $v_i \in P$, $d_i \in \mathbb{Z}_{\geq 1}$, and $U = (u_1, \ldots, u_n)$, where $u_i : \mathsf{X} \to \mathbb{R}$ is the payoff function of player v_i, $i \in \{1, \ldots, n\}$. The game $\mathbf{G}$ is called a *zero-sum game* if $\sum_{i=1}^{n} u_i = 0$. An outcome $x^* \in \mathsf{X}$ is a (pure) *Nash equilibrium* of $\mathbf{G}$ if for all $i \in \{1, \ldots, n\}$ and all $x_i \in \mathsf{X}_i$,

$$u_i(x_i^*, x_{-i}^*) \geq u_i(x_i, x_{-i}^*),$$

where x_{-i} denotes the actions of all players other than v_i. In this paper, we focus on a class of two-player zero-sum games which have at least one pure Nash equilibrium as the next result states.

Theorem 2.5 (Minmax theorem): *Let $\mathsf{X}_1 \subset \mathbb{R}^{d_1}$ and $\mathsf{X}_2 \subset \mathbb{R}^{d_2}$, $d_1, d_2 \in \mathbb{Z}_{\geq 1}$, be nonempty, compact, and convex. If $u : \mathsf{X}_1 \times \mathsf{X}_2 \to \mathbb{R}$ is continuous and the sets $\{x' \in \mathsf{X}_1 \mid u(x', y) \geq \alpha\}$ and $\{y' \in \mathsf{X}_2 \mid u(x, y') \leq \alpha\}$ are convex for all $x \in \mathsf{X}_1$, $y \in \mathsf{X}_2$, and $\alpha \in \mathbb{R}$, then*

$$\max_x \min_y u(x, y) = \min_y \max_x u(x, y).$$

Theorem 2.5 implies that the game $\mathbf{G} = (\{v_1, v_2\}, \mathsf{X}_1 \times \mathsf{X}_2, (u, -u))$ has a pure Nash equilibrium.

3 Problem statement

Consider two networks Σ_1 and Σ_2 composed of agents $\{v_1, \ldots, v_{n_1}\}$ and agents $\{w_1, \ldots, w_{n_2}\}$, respectively. Throughout this paper, Σ_1 and Σ_2 are either connected undirected graphs, c.f. Section 4, or strongly connected weight-balanced digraphs, c.f. Section 5. Since the latter case includes the first one, throughout this section, we assume the latter. The state of Σ_1, denoted x_1, belongs to $\mathsf{X}_1 \subset \mathbb{R}^{d_1}$, $d_1 \in \mathbb{Z}_{\geq 1}$. Likewise, the state of Σ_2, denoted x_2, belongs to $\mathsf{X}_2 \subset \mathbb{R}^{d_2}$, $d_2 \in \mathbb{Z}_{\geq 1}$. In this paper, we do not get into the details of what these states represent (as a particular case, the network state could correspond to the collection of the states of agents in it). In addition, each agent v_i in Σ_1 has an estimate $x_1^i \in \mathbb{R}^{d_1}$ of what the network state is, which may differ from the actual value x_1. Similarly, each agent w_j in Σ_2 has an estimate $x_2^j \in \mathbb{R}^{d_2}$ of what the network state is. Within each network, neighboring agents can share their estimates. Networks can also obtain information about each other. This is modeled by means of a bipartite directed graph Σ_{eng}, called *engagement* graph, with disjoint vertex sets $\{v_1, \ldots, v_{n_1}\}$ and $\{w_1, \ldots, w_{n_2}\}$, where every agent has at least one out-neighbor. According to this model, an agent in Σ_1 obtains information from its out-neighbors in Σ_{eng} about their estimates of the state of Σ_2, and vice versa.

Figure 1 illustrates this concept.

For each $i \in \{1, \ldots, n_1\}$, let $f_1^i : \mathsf{X}_1 \times \mathsf{X}_2 \to \mathbb{R}$ be a locally Lipschitz concave-convex function only available to agent $v_i \in \Sigma_1$. Similarly, let $f_2^j : \mathsf{X}_1 \times \mathsf{X}_2 \to \mathbb{R}$ be a locally Lipschitz concave-convex function only available to agent $w_j \in \Sigma_2$, $j \in \{1, \ldots, n_2\}$. The networks Σ_1 and Σ_2 are engaged in a zero-sum game with payoff function $U : \mathsf{X}_1 \times \mathsf{X}_2 \to \mathbb{R}$

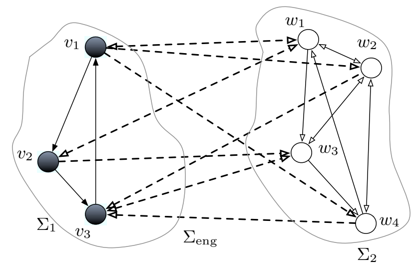


Fig. 1. Networks Σ_1 and Σ_2 engaged in a strategic scenario. Both networks are strongly connected and weight-balanced, with weights of 1 on each edge. Edges which correspond to Σ_{eng} are dashed.

$$U(x_1, x_2) = \sum_{i=1}^{n_1} f_1^i(x_1, x_2) = \sum_{j=1}^{n_2} f_2^j(x_1, x_2), \qquad (3)$$

where Σ_1 wishes to maximize U, while Σ_2 wishes to minimize it. The objective of the networks is therefore to settle upon a Nash equilibrium, i.e., to solve the following maxmin problem

$$\max_{x_1 \in \mathsf{X}_1} \min_{x_2 \in \mathsf{X}_2} U(x_1, x_2). \qquad (4)$$

We refer to the this zero-sum game as the 2-*network zero-sum game* and denote it by $\mathbf{G}_{\text{adv-net}} = (\Sigma_1, \Sigma_2, \Sigma_{\text{eng}}, U)$. We assume that $\mathsf{X}_1 \subset \mathbb{R}^{d_1}$ and $\mathsf{X}_2 \subset \mathbb{R}^{d_2}$ are compact convex. For convenience, let $\boldsymbol{x}_1 = (x_1^1, \ldots, x_1^{n_1})^T$ and $\boldsymbol{x}_2 = (x_2^1, \ldots, x_2^{n_2})^T$ denote vector of agent estimates about the state of the respective networks.

Remark 3.1 (Power allocation in communication channels in the presence of adversaries): Here we present an example from communications inspired by (Boyd and Vandenberghe, 2004, Section 5.5.3). Consider n Gaussian communication channels, each with signal power $p_i \in \mathbb{R}_{\geq 0}$ and noise power $\eta_i \in \mathbb{R}_{\geq 0}$, for $i \in \{1, \ldots, n\}$. The capacity of each channel is proportional to $\log(1 + \beta p_i/(\sigma_i + \eta_i))$, where $\beta \in \mathbb{R}_{>0}$ and $\sigma_i > 0$ is the receiver noise. Note that capacity is concave in p_i and convex in η_i. Both signal and noise powers must satisfy a budget constraint, i.e., $\sum_{i=1}^{n} p_i = P$ and $\sum_{i=1}^{n} \eta_i = C$, for some given $P, C \in \mathbb{R}_{>0}$. Two networks of n agents are involved in this scenario, one, Σ_1, selecting signal powers to maximize capacity, the other one, Σ_2, selecting noise powers to minimize it. The network Σ_1 has decided that m_1 channels will have signal power x_1, while $n - 1 - m_1$ will have signal power x_2. The remaining nth channel has its power determined to satisfy the budget constraint, i.e., $P - m_1 x_1 - (n - 1 - m_1) x_2$. Likewise, the network Σ_2 does something similar with m_2 channels with noise power y_1, $n - 1 - m_2$ channels with noise power y_2, and one last channel with noise power $C - m_2 y_1 - (n - 1 - m_2) y_2$. Each network is aware of the partition made by the other one. The individual

objective function of the two agents (one from Σ_1, the other from Σ_2) making decisions on the power levels of the ith channel is the channel capacity itself. For $i \in \{1, \dots, n-1\}$, this takes the form

$$f^i(x, y) = \log\Big(1 + \frac{\beta x_a}{\sigma_i + y_b}\Big),$$

for some $a, b \in \{1, 2\}$. Here $x = (x_1, x_2)$ and $y = (y_1, y_2)$. For $i = n$, it takes instead the form

$$f^n(x, y) = \log\Big(1 + \frac{\beta(P - m_1 x_1 - (n-1-m_1)x_2)}{\sigma_n + C - m_2 y_1 - (n-1-m_2)y_2}\Big).$$

Note that $\sum_{i=1}^n f^i(x, y)$ is the total capacity of the n communication channels. •

3.1 Reformulation of the 2-network zero-sum game

In this section, we describe how agents in each network use the information obtained from their neighbors to compute the value of their own objective functions. Based on these estimates, we introduce a reformulation of the $\mathbf{G}_{\text{adv-net}} = (\Sigma_1, \Sigma_2, \Sigma_{\text{eng}}, U)$ which is instrumental for establishing some of our results.

Each agent in Σ_1 has a locally Lipschitz, concave-convex function $\tilde{f}_1^i : \mathbb{R}^{d_1} \times \mathbb{R}^{d_2 n_2} \to \mathbb{R}$ with the properties:

(Extension of own payoff function): for any $x_1 \in \mathbb{R}^{d_1}$, $x_2 \in \mathbb{R}^{d_2}$,

$$\tilde{f}_1^i(x_1, \mathbf{1}_{n_2} \otimes x_2) = f_1^i(x_1, x_2). \tag{5a}$$

(Distributed over Σ_{eng}): there exists $\mathfrak{f}_1^i : \mathbb{R}^{d_1} \times \mathbb{R}^{d_2 |\mathcal{N}^{\text{in}}_{\Sigma_{\text{eng}}}(v_i)|} \to \mathbb{R}$ such that, for any $x_1 \in \mathbb{R}^{d_1}$ $\boldsymbol{x}_2 \in \mathbb{R}^{d_2 n_2}$,

$$\tilde{f}_1^i(x_1, \boldsymbol{x}_2) = \mathfrak{f}_1^i(x_1, \pi_1^i(\boldsymbol{x}_2)), \tag{5b}$$

with $\pi_1^i : \mathbb{R}^{d_2 n_2} \to \mathbb{R}^{d_2 |\mathcal{N}^{\text{out}}_{\Sigma_{\text{eng}}}(v_i)|}$ the projection of $\boldsymbol{x}_2$ to the values received by v_i from its out-neighbors in Σ_{eng}.

Equation (5a) states the fact that, when the estimates of all neighbors of an agent in the opponent's network agree, its evaluation should coincide with this estimate. Equation (5b) states the fact that agents can only use the information received from their neighbors in the interaction topology to compute their new estimates.

Each agent in Σ_2 has a function $\tilde{f}_2^j : \mathbb{R}^{d_1 n_1} \times \mathbb{R}^{d_2} \to \mathbb{R}$ with similar properties. The collective payoff functions of the two networks are

$$\tilde{U}_1(\boldsymbol{x}_1, \boldsymbol{x}_2) = \sum_{i=1}^{n_1} \tilde{f}_1^i(x_1^i, \boldsymbol{x}_2), \tag{6a}$$

$$\tilde{U}_2(\boldsymbol{x}_1, \boldsymbol{x}_2) = \sum_{j=1}^{n_2} \tilde{f}_2^j(\boldsymbol{x}_1, x_2^j). \tag{6b}$$

In general, the functions $\tilde{U}_1$ and $\tilde{U}_2$ need not be the same. However, $\tilde{U}_1(\mathbf{1}_{n_1} \otimes x_1, \mathbf{1}_{n_1} \otimes x_2) = \tilde{U}_2(\mathbf{1}_{n_1} \otimes x_1, \mathbf{1}_{n_1} \otimes x_2)$, for any $x_1 \in \mathbb{R}^{d_1}$, $x_2 \in \mathbb{R}^{d_2}$. When both functions coincide, the next result shows that the original game can be lifted to a (constrained) zero-sum game.

Lemma 3.2 (Reformulation of the 2-network zero-sum game): *Assume that the individual payoff functions $\{\tilde{f}_1^i\}_{i=1}^{n_1}$, $\{\tilde{f}_2^j\}_{j=1}^{n_2}$ satisfying (5) are such that the network payoff functions defined in (6) satisfy $\tilde{U}_1 = \tilde{U}_2$, and let $\tilde{U}$ denote this common function. Then, the problem (4) on $\mathbb{R}^{d_1} \times \mathbb{R}^{d_2}$ is equivalent to the following problem on $\mathbb{R}^{n_1 d_1} \times \mathbb{R}^{n_2 d_2}$,*

$$\begin{aligned} &\max_{\boldsymbol{x}_1 \in \mathsf{X}_1^{n_1}} \min_{\boldsymbol{x}_2 \in \mathsf{X}_2^{n_2}} \tilde{U}(\boldsymbol{x}_1, \boldsymbol{x}_2), \\ &\qquad \textit{subject to} \quad \mathbf{L}_1 \boldsymbol{x}_1 = \mathbf{0}_{n_1 d_1}, \quad \mathbf{L}_2 \boldsymbol{x}_2 = \mathbf{0}_{n_2 d_2}, \end{aligned} \tag{7}$$

with $\mathbf{L}_\ell = \mathsf{L}_\ell \otimes \mathsf{I}_{d_\ell}$ and L_ℓ the Laplacian of Σ_ℓ, $\ell \in \{1, 2\}$.

Proof. The proof follows by noting that (i) $\tilde{U}(\mathbf{1}_{n_1} \otimes x_1, \mathbf{1}_{n_2} \otimes x_2) = U(x_1, x_2)$ for all $x_1 \in \mathbb{R}^{d_1}$ and $x_2 \in \mathbb{R}^{d_2}$ and (ii) since $\mathcal{G}_1$ and $\mathcal{G}_2$ are strongly connected, $\mathbf{L}_1 \boldsymbol{x}_1 = \mathbf{0}_{n_1 d_1}$ and $\mathbf{L}_2 \boldsymbol{x}_2 = \mathbf{0}_{n_2 d_2}$ iff $\boldsymbol{x}_1 = \mathbf{1}_{n_1} \otimes x_1$ and $\boldsymbol{x}_2 = \mathbf{1}_{n_2} \otimes x_2$ for some $x_1 \in \mathbb{R}^{d_1}$ and $x_2 \in \mathbb{R}^{d_2}$. □

Remark 3.3 (Restrictions on extensions): The assumption of Lemma 3.2 does not hold in general for all sets of extensions satisfying (5a) and (5b). If the interaction topology is one-to-one (i.e., both networks have the same number of agents, the interaction topology is undirected, and each agent in the first network obtains information only from one agent in the opposing network), the natural extensions satisfy the assumption. Example 5.5 later provides yet another instance of a different nature. In general, determining if it is always possible to choose the extensions in such a way that the assumption holds is an open problem. •

We denote by $\tilde{\mathbf{G}}_{\text{adv-net}} = (\Sigma_1, \Sigma_2, \Sigma_{\text{eng}}, \tilde{U})$ the constrained zero-sum game defined by (7) and refer to this situation by saying that $\mathbf{G}_{\text{adv-net}}$ can be lifted to $\tilde{\mathbf{G}}_{\text{adv-net}}$. Our objective is to design a coordination algorithm that is implementable with the information that agents in Σ_1 and Σ_2 possess and leads them to find a Nash equilibrium of $\tilde{\mathbf{G}}_{\text{adv-net}}$, which corresponds to a Nash equilibrium of $\mathbf{G}_{\text{adv-net}}$ by Lemma 3.2. Achieving this goal, however, is nontrivial because individual agents, not networks themselves, are the decision makers. From the point of view of agents in each network, the objective is to agree on the states of both their own network and the other network, and that the resulting states correspond to a Nash equilibrium of $\mathbf{G}_{\text{adv-net}}$.

The function $\tilde{U}$ is locally Lipschitz and concave-convex. Moreover, from Lemma 2.2, the elements of $\partial_{\boldsymbol{x}_1} \tilde{U}(\boldsymbol{x}_1, \boldsymbol{x}_2)$ are of the form

$$\tilde{g}_{(\boldsymbol{x}_1, \boldsymbol{x}_2)} = (\tilde{g}^1_{(x_1^1, \boldsymbol{x}_2)}, \dots, \tilde{g}^n_{(x_1^n, \boldsymbol{x}_2)}) \in \partial_{\boldsymbol{x}_1} \tilde{U}(\boldsymbol{x}_1, \boldsymbol{x}_2),$$

where $\tilde{g}^i_{(x_1^i,\boldsymbol{x}_2)} \in \partial_{x_1}\tilde{f}_1^i(x_1^i,\boldsymbol{x}_2)$, for $i\in\{1,\dots,n_1\}$. Note that, because of (5b), we have $\partial_{x_1}\tilde{f}_1^i(x_1^i,\mathbf{1}_{n_2}\otimes x_2) = \partial_{x_1}f_1^i(x_1^i,x_2)$. A similar reasoning can be followed to describe the elements of $\partial_{\boldsymbol{x}_2}\tilde{U}(\boldsymbol{x}_1,\boldsymbol{x}_2)$. Next, we present a characterization of the Nash equilibria of $\tilde{\mathbf{G}}_{\text{adv-net}}$, instrumental for proving some of our upcoming results.

Proposition 3.4 (Characterization of the Nash equilibria of $\tilde{\mathbf{G}}_{\text{adv-net}}$): *For Σ_1, Σ_2 strongly connected and weight-balanced, define F_1 and F_2 by*

$$F_1(\boldsymbol{x}_1,\boldsymbol{z}_1,\boldsymbol{x}_2) = -\tilde{U}(\boldsymbol{x}_1,\boldsymbol{x}_2) + \boldsymbol{x}_1^T\mathbf{L}_1\boldsymbol{z}_1 + \frac{1}{2}\boldsymbol{x}_1^T\mathbf{L}_1\boldsymbol{x}_1,$$
$$F_2(\boldsymbol{x}_2,\boldsymbol{z}_2,\boldsymbol{x}_1) = \tilde{U}(\boldsymbol{x}_1,\boldsymbol{x}_2) + \boldsymbol{x}_2^T\mathbf{L}_2\boldsymbol{z}_2 + \frac{1}{2}\boldsymbol{x}_2^T\mathbf{L}_2\boldsymbol{x}_2.$$

Then, F_1 and F_2 are convex in their first argument, linear in their second one, and concave in their third one. Moreover, assume $(\boldsymbol{x}_1^*,\boldsymbol{z}_1^*,\boldsymbol{x}_2^*,\boldsymbol{z}_2^*)$ satisfies the following saddle property *for (F_1,F_2): $(\boldsymbol{x}_1^*,\boldsymbol{z}_1^*)$ is a saddle point of $(\boldsymbol{x}_1,\boldsymbol{z}_1)\mapsto F_1(\boldsymbol{x}_1,\boldsymbol{z}_1,\boldsymbol{x}_2^*)$ and $(\boldsymbol{x}_2^*,\boldsymbol{z}_2^*)$ is a saddle point of $(\boldsymbol{x}_2,\boldsymbol{z}_2)\mapsto F_2(\boldsymbol{x}_2,\boldsymbol{z}_2,\boldsymbol{x}_1^*)$. Then,*

(i) $(\boldsymbol{x}_1^*,\boldsymbol{z}_1^*+\mathbf{1}_{n_1}\otimes a_1,\boldsymbol{x}_2^*,\boldsymbol{z}_2^*+\mathbf{1}_{n_2}\otimes a_2)$ satisfies the saddle property for (F_1,F_2) for any $a_1\in\mathbb{R}^{d_1}$, $a_2\in\mathbb{R}^{d_2}$, and

(ii) $(\boldsymbol{x}_1^*,\boldsymbol{x}_2^*)$ is a Nash equilibrium of $\tilde{\mathbf{G}}_{\text{adv-net}}$.

Furthermore,

(iii) if $(\boldsymbol{x}_1^*,\boldsymbol{x}_2^*)$ is a Nash equilibrium of $\tilde{\mathbf{G}}_{\text{adv-net}}$ then there exists $\boldsymbol{z}_1^*,\boldsymbol{z}_2^*$ such that $(\boldsymbol{x}_1^*,\boldsymbol{z}_1^*,\boldsymbol{x}_2^*,\boldsymbol{z}_2^*)$ satisfies the saddle property for (F_1,F_2).

Proof. The statement (i) is immediate. For (ii), since $(\boldsymbol{x}_1^*,\boldsymbol{z}_1^*,\boldsymbol{x}_2^*,\boldsymbol{z}_2^*)$ satisfies the saddle property, and the networks are strongly connected and weight-balanced, we have $\boldsymbol{x}_1^* = \mathbf{1}_{n_1}\otimes x_1^*$, $x_1^*\in\mathbb{R}^{d_1}$, $\boldsymbol{x}_2^*=\mathbf{1}_{n_2}\otimes x_2^*$, $x_2^*\in\mathbb{R}^{d_2}$, $\mathbf{L}_1\boldsymbol{z}_1^* \in -\partial_{\boldsymbol{x}_1}\tilde{U}(\boldsymbol{x}_1^*,\boldsymbol{x}_2^*)$, and $\mathbf{L}_2\boldsymbol{z}_2^* \in \partial_{\boldsymbol{x}_2}\tilde{U}(\boldsymbol{x}_1^*,\boldsymbol{x}_2^*)$. Thus there exist $g^i_{1,(x_1^*,x_2^*)} \in \partial_{x_1}f_1^i(x_1^*,x_2^*)$, $i\in\{1,\dots,n_1\}$, and $g^j_{2,(x_1^*,x_2^*)} \in -\partial_{x_2}f_2^j(x_1^*,x_2^*)$, $j\in\{1,\dots,n_2\}$, such that

$$\mathbf{L}_1\boldsymbol{z}_1^* = (g^1_{1,(x_1^*,x_2^*)},,\dots,g^n_{1,(x_1^*,x_2^*)}))^T, \text{ and}$$
$$\mathbf{L}_2\boldsymbol{z}_2^* = (g^1_{2,(x_1^*,x_2^*)},,\dots,g^n_{2,(x_1^*,x_2^*)}))^T.$$

Noting that, for $\ell\in\{1,2\}$, $(\mathbf{1}_{n_\ell}^T\otimes \mathsf{I}_{d_\ell})\mathbf{L}_\ell = (\mathbf{1}_{n_\ell}^T\otimes \mathsf{I}_{d_\ell})(\mathsf{L}_\ell\otimes\mathsf{I}_{d_\ell}) = \mathbf{1}_{n_\ell}^T\mathsf{L}\otimes\mathsf{I}_{d_\ell} = \mathbf{0}_{d_\ell\times d_\ell n_\ell}$, we deduce that $\sum_{i=1}^{n_1} g^i_{1,(x_1^*,x_2^*)} = \mathbf{0}_{d_1}$ and $\sum_{j=1}^{n_2} g^j_{2,(x_1^*,x_2^*)} = \mathbf{0}_{d_2}$, i.e., $(\boldsymbol{x}_1^*,\boldsymbol{x}_2^*)$ is a Nash equilibrium. Finally for proving (iii), note that $\boldsymbol{x}_1^*=\mathbf{1}_{n_1}\otimes x_1^*$ and $\boldsymbol{x}_2^*=\mathbf{1}_{n_2}\otimes x_2^*$. The result follows then from the fact that $0\in\partial_{x_1}U(x_1^*,x_2^*)$ and $0\in\partial_{x_2}U(x_1^*,x_2^*)$ implies that there exists $\boldsymbol{z}_1^*\in\mathbb{R}^{n_1d_1}$ and $\boldsymbol{z}_2^*\in\mathbb{R}^{n_2d_2}$ with $\mathbf{L}_1\boldsymbol{z}_1^*\in\partial_{\boldsymbol{x}_1}\tilde{U}(\boldsymbol{x}_1^*,\boldsymbol{x}_2^*)$ and $\mathbf{L}_2\boldsymbol{z}_2^*\in -\partial_{\boldsymbol{x}_2}\tilde{U}(\boldsymbol{x}_1^*,\boldsymbol{x}_2^*)$. □

4 Distributed convergence to Nash equilibria for undirected topologies

In this section, we introduce a distributed dynamics which solves (7) when Σ_1 and Σ_2 are undirected. In particular, we design gradient dynamics to find points with the saddle property for (F_1,F_2) prescribed by Proposition 3.4. Consider the set-valued dynamics $\Psi_{\text{Nash-undir}} : (\mathbb{R}^{d_1n_1})^2\times(\mathbb{R}^{d_2n_2})^2 \rightrightarrows (\mathbb{R}^{d_1n_1})^2\times(\mathbb{R}^{d_2n_2})^2$,

$$\dot{\boldsymbol{x}}_1 + \mathbf{L}_1\boldsymbol{x}_1 + \mathbf{L}_1\boldsymbol{z}_1 \in \partial_{\boldsymbol{x}_1}\tilde{U}(\boldsymbol{x}_1,\boldsymbol{x}_2), \tag{8a}$$
$$\dot{\boldsymbol{z}}_1 = \mathbf{L}_1\boldsymbol{x}_1, \tag{8b}$$
$$\dot{\boldsymbol{x}}_2 + \mathbf{L}_2\boldsymbol{x}_2 + \mathbf{L}_2\boldsymbol{z}_2 \in -\partial_{\boldsymbol{x}_2}\tilde{U}(\boldsymbol{x}_1,\boldsymbol{x}_2), \tag{8c}$$
$$\dot{\boldsymbol{z}}_2 = \mathbf{L}_2\boldsymbol{x}_2, \tag{8d}$$

where $\boldsymbol{x}_\ell,\boldsymbol{z}_\ell\in\mathbb{R}^{n_\ell d_\ell}$, $\ell\in\{1,2\}$. Note that (8a)-(8b) and (8c)-(8d) correspond to saddle-point dynamics of F_1 in $(\boldsymbol{x}_1,\boldsymbol{z}_1)$ and F_2 in $(\boldsymbol{x}_2,\boldsymbol{z}_2)$, respectively. Local solutions to this dynamics exist by virtue of Lemmas 2.1 and 2.3. We characterize next its asymptotic convergence properties.

Theorem 4.1 (Distributed convergence to Nash equilibria for undirected networks): *Consider the zero-sum game $\mathbf{G}_{\text{adv-net}} = (\Sigma_1,\Sigma_2,\Sigma_{\text{eng}},U)$, with Σ_1 and Σ_2 connected undirected graphs, $\mathsf{X}_1\subset\mathbb{R}^{d_1}$, $\mathsf{X}_2\subset\mathbb{R}^{d_2}$ compact and convex, and $U:\mathsf{X}_1\times\mathsf{X}_2\to\mathbb{R}$ strictly concave-convex and locally Lipschitz. Assume $\mathbf{G}_{\text{adv-net}}$ can be lifted to $\tilde{\mathbf{G}}_{\text{adv-net}}$. Then, the projection onto the first and third components of the solutions of (8) asymptotically converge to agreement on the Nash equilibrium of $\mathbf{G}_{\text{adv-net}}$.*

Proof. Throughout this proof, since property (5b) holds, without loss of generality and for simplicity of notation, we assume that agents in Σ_1 have access to $\boldsymbol{x}_2$ and, similarly, agents in Σ_2 have access to $\boldsymbol{x}_1$. By Theorem 2.5, a solution to (6) exists. By the strict concavity-convexity properties, this solution is, in fact, unique. Let us denote this solution by $\boldsymbol{x}_1^*=\mathbf{1}_{n_1}\otimes x_1^*$ and $\boldsymbol{x}_2^*=\mathbf{1}_{n_2}\otimes x_2^*$. By Proposition 3.4(iii), there exists $\boldsymbol{z}_1^*$ and $\boldsymbol{z}_2^*$ such that $(\boldsymbol{x}_1^*,\boldsymbol{z}_1^*,\boldsymbol{x}_2^*,\boldsymbol{z}_2^*)\in\mathrm{Eq}(\Psi_{\text{Nash-undir}})$. First, note that given any initial condition $(\boldsymbol{x}_1^0,\boldsymbol{z}_1^0,\boldsymbol{x}_2^0,\boldsymbol{z}_2^0)\in(\mathbb{R}^{n_1d_1})^2\times(\mathbb{R}^{n_2d_2})^2$, the set

$$W_{\boldsymbol{z}_1^0,\boldsymbol{z}_2^0} = \{(\boldsymbol{x}_1,\boldsymbol{z}_1,\boldsymbol{x}_2,\boldsymbol{z}_2) \mid (\mathbf{1}_{n_\ell}^T\otimes\mathsf{I}_{d_\ell})\boldsymbol{z}_\ell = (\mathbf{1}_{n_\ell}^T\otimes\mathsf{I}_{d_\ell})\boldsymbol{z}_\ell^0,\ \ell\in\{1,2\}\} \tag{9}$$

is strongly positively invariant under (8). Consider the function $V:(\mathbb{R}^{d_1n_1})^2\times(\mathbb{R}^{d_2n_2})^2\to\mathbb{R}_{\geq 0}$ defined by

$$\begin{aligned}V(\boldsymbol{x}_1,\boldsymbol{z}_1,\boldsymbol{x}_2,\boldsymbol{z}_2) &= \frac{1}{2}(\boldsymbol{x}_1-\boldsymbol{x}_1^*)^T(\boldsymbol{x}_1-\boldsymbol{x}_1^*) + \frac{1}{2}(\boldsymbol{z}_1-\boldsymbol{z}_1^*)^T(\boldsymbol{z}_1-\boldsymbol{z}_1^*)\\ &\quad + \frac{1}{2}(\boldsymbol{x}_2-\boldsymbol{x}_2^*)^T(\boldsymbol{x}_2-\boldsymbol{x}_2^*) + \frac{1}{2}(\boldsymbol{z}_2-\boldsymbol{z}_2^*)^T(\boldsymbol{z}_2-\boldsymbol{z}_2^*).\end{aligned}$$

The function V is smooth. Next, we examine its set-valued Lie derivative along $\Psi_{\text{Nash-undir}}$. Let $\xi \in \widetilde{\mathcal{L}}_{\Psi_{\text{Nash-undir}}}V(\boldsymbol{x}_1,\boldsymbol{z}_1,\boldsymbol{x}_2,\boldsymbol{z}_2)$. By definition, there exists $v \in \Psi_{\text{Nash-undir}}(\boldsymbol{x}_1,\boldsymbol{z}_1,\boldsymbol{x}_2,\boldsymbol{z}_2)$, given by

$$\begin{aligned} v = (&-\mathbf{L}_1\boldsymbol{x}_1 - \mathbf{L}_1\boldsymbol{z}_1 + g_{1,(\boldsymbol{x}_1,\boldsymbol{x}_2)}, \\ &-\mathbf{L}_2\boldsymbol{x}_2 - \mathbf{L}_2\boldsymbol{z}_2 - g_{2,(\boldsymbol{x}_1,\boldsymbol{x}_2)}, \mathbf{L}_1\boldsymbol{x}_1, \mathbf{L}_2\boldsymbol{x}_2), \end{aligned}$$

where $g_{1,(\boldsymbol{x}_1,\boldsymbol{x}_2)} \in \partial_{\boldsymbol{x}_1}U(\boldsymbol{x}_1,\boldsymbol{x}_2)$ and $g_{2,(\boldsymbol{x}_1,\boldsymbol{x}_2)} \in \partial_{\boldsymbol{x}_2}U(\boldsymbol{x}_1,\boldsymbol{x}_2)$, such that

$$\begin{aligned} \xi &= v \cdot \nabla V(\boldsymbol{x}_1,\boldsymbol{z}_1,\boldsymbol{x}_2,\boldsymbol{z}_2) \\ &= (\boldsymbol{x}_1 - \boldsymbol{x}_1^*)^T(-\mathbf{L}_1\boldsymbol{x}_1 - \mathbf{L}_1\boldsymbol{z}_1 + g_{1,(\boldsymbol{x}_1,\boldsymbol{x}_2)}) \\ &+ (\boldsymbol{x}_2 - \boldsymbol{x}_2^*)^T(-\mathbf{L}_2\boldsymbol{x}_2 - \mathbf{L}_2\boldsymbol{z}_2 - g_{2,(\boldsymbol{x}_1,\boldsymbol{x}_2)}) \\ &\quad + (\boldsymbol{z}_1 - \boldsymbol{z}_1^*)^T\mathbf{L}_1\boldsymbol{x}_1 + (\boldsymbol{z}_2 - \boldsymbol{z}_2^*)^T\mathbf{L}_2\boldsymbol{x}_2. \end{aligned}$$

Note that $-\mathbf{L}_1\boldsymbol{x}_1 - \mathbf{L}_1\boldsymbol{z}_1 + g_{1,(\boldsymbol{x}_1,\boldsymbol{x}_2)} \in -\partial_{\boldsymbol{x}_1}F_1(\boldsymbol{x}_1,\boldsymbol{z}_1,\boldsymbol{x}_2)$, $\mathbf{L}_1\boldsymbol{x}_1 \in \partial_{\boldsymbol{z}_1}F_1(\boldsymbol{x}_1,\boldsymbol{z}_1,\boldsymbol{x}_2)$, $-\mathbf{L}_2\boldsymbol{x}_2 - \mathbf{L}_2\boldsymbol{z}_2 - g_{2,(\boldsymbol{x}_1,\boldsymbol{x}_2)} \in -\partial_{\boldsymbol{x}_2}F_2(\boldsymbol{x}_2,\boldsymbol{z}_2,\boldsymbol{x}_2)$, and $\mathbf{L}_2\boldsymbol{x}_2 \in \partial_{\boldsymbol{z}_2}F_2(\boldsymbol{x}_2,\boldsymbol{z}_2,\boldsymbol{x}_1)$. Using the first-order convexity property of F_1 and F_2 in their first two arguments, one gets

$$\begin{aligned} \xi \le\, & F_1(\boldsymbol{x}_1^*,\boldsymbol{z}_1,\boldsymbol{x}_2) - F_1(\boldsymbol{x}_1,\boldsymbol{z}_1,\boldsymbol{x}_2) + F_2(\boldsymbol{x}_2^*,\boldsymbol{z}_2,\boldsymbol{x}_1) \\ & - F_2(\boldsymbol{x}_2,\boldsymbol{z}_2,\boldsymbol{x}_1) + F_1(\boldsymbol{x}_1,\boldsymbol{z}_1,\boldsymbol{x}_2) - F_1(\boldsymbol{x}_1,\boldsymbol{z}_1^*,\boldsymbol{x}_2) \\ & + F_2(\boldsymbol{x}_2,\boldsymbol{z}_2,\boldsymbol{x}_1) - F_2(\boldsymbol{x}_2,\boldsymbol{z}_2^*,\boldsymbol{x}_1). \end{aligned}$$

Expanding each term and using the fact that $(\boldsymbol{x}_1^*,\boldsymbol{z}_1^*,\boldsymbol{x}_2^*,\boldsymbol{z}_2^*) \in \mathrm{Eq}(\Psi_{\text{Nash-undir}})$, we simplify this inequality as

$$\begin{aligned} \xi \le\, & -\tilde{U}(\boldsymbol{x}_1^*,\boldsymbol{x}_2) + \tilde{U}(\boldsymbol{x}_1,\boldsymbol{x}_2^*) - \boldsymbol{z}_1^*\mathbf{L}_1\boldsymbol{x}_1 \\ & - \frac{1}{2}\boldsymbol{x}_1\mathbf{L}_1\boldsymbol{x}_1 - \boldsymbol{z}_2^*\mathbf{L}_2\boldsymbol{x}_2 - \frac{1}{2}\boldsymbol{x}_2\mathbf{L}_2\boldsymbol{x}_2. \end{aligned}$$

By rearranging, we thus have

$$\xi \le -F_2(\boldsymbol{x}_2,\boldsymbol{z}_2^*,\boldsymbol{x}_1^*) - F_1(\boldsymbol{x}_1,\boldsymbol{z}_1^*,\boldsymbol{x}_2^*).$$

Next, since $F_2(\boldsymbol{x}_1^*,\boldsymbol{z}_2^*,\boldsymbol{x}_2^*) + F_1(\boldsymbol{x}_2^*,\boldsymbol{z}_2^*,\boldsymbol{x}_1^*) = 0$, we have

$$\begin{aligned} \xi \le\, & F_1(\boldsymbol{x}_1^*,\boldsymbol{z}_1^*,\boldsymbol{x}_2^*) - F_1(\boldsymbol{x}_1,\boldsymbol{z}_1^*,\boldsymbol{x}_2^*) \\ & + F_2(\boldsymbol{x}_2^*,\boldsymbol{z}_2^*,\boldsymbol{x}_1^*) - F_2(\boldsymbol{x}_2,\boldsymbol{z}_2^*,\boldsymbol{x}_1^*), \end{aligned}$$

yielding that $\xi \le 0$. As a result,

$$\max \widetilde{\mathcal{L}}_{\Psi_{\text{Nash-undir}}}V(\boldsymbol{x}_1,\boldsymbol{z}_1,\boldsymbol{x}_2,\boldsymbol{z}_2) \le 0.$$

As a by-product, we conclude that the trajectories of (8) are bounded. By virtue of the set-valued version of the LaSalle Invariance Principle, cf. Theorem 2.4, any trajectory of (8) starting from an initial condition $(\boldsymbol{x}_1^0,\boldsymbol{z}_1^0,\boldsymbol{x}_2^0,\boldsymbol{z}_2^0)$ converges to the largest positively invariant set M in $S_{\Psi_{\text{Nash-undir}},V} \cap V^{-1}(\le V(\boldsymbol{x}_1^0,\boldsymbol{z}_1^0,\boldsymbol{x}_2^0,\boldsymbol{z}_2^0))$. Let $(\boldsymbol{x}_1,\boldsymbol{z}_1,\boldsymbol{x}_2,\boldsymbol{z}_2) \in M$. Because $M \subset S_{\Psi_{\text{Nash-undir}},V}$, then $F_1(\boldsymbol{x}_1^*,\boldsymbol{z}_1^*,\boldsymbol{x}_2^*) - F_1(\boldsymbol{x}_1,\boldsymbol{z}_1^*,\boldsymbol{x}_2^*) = 0$, i.e.,

$$-\tilde{U}(\boldsymbol{x}_1^*,\boldsymbol{x}_2^*) + \tilde{U}(\boldsymbol{x}_1,\boldsymbol{x}_2^*) - \boldsymbol{x}_1^T\mathbf{L}_1\boldsymbol{z}_1^* - \frac{1}{2}\boldsymbol{x}_1^T\mathbf{L}_1\boldsymbol{x}_1 = 0. \tag{10}$$

Define now $G_1 : \mathbb{R}^{n_1d_1} \times \mathbb{R}^{n_1d_1} \times \mathbb{R}^{n_2d_2} \to \mathbb{R}$ by $G_1(\boldsymbol{x}_1,\boldsymbol{z}_1,\boldsymbol{x}_2) = F_1(\boldsymbol{x}_1,\boldsymbol{z}_1,\boldsymbol{x}_2) - \frac{1}{2}\boldsymbol{x}_1^T\mathbf{L}_1\boldsymbol{x}_1$. G_1 is convex in its first argument and linear in its second. Furthermore, for fixed $\boldsymbol{x}_2$, the map $(\boldsymbol{x}_1,\boldsymbol{z}_1) \mapsto G_1(\boldsymbol{x}_1,\boldsymbol{z}_1,\boldsymbol{x}_2)$ has the same saddle points as $(\boldsymbol{x}_1,\boldsymbol{z}_1) \mapsto F_1(\boldsymbol{x}_1,\boldsymbol{z}_1,\boldsymbol{x}_2)$. As a result, $G_1(\boldsymbol{x}_1^*,\boldsymbol{z}_1^*,\boldsymbol{x}_2^*) - G_1(\boldsymbol{x}_1,\boldsymbol{z}_1^*,\boldsymbol{x}_2^*) \le 0$, or equivalently, $-\tilde{U}(\boldsymbol{x}_1^*,\boldsymbol{x}_2^*) + \tilde{U}(\boldsymbol{x}_1,\boldsymbol{x}_2^*) - \boldsymbol{x}_1^T\mathbf{L}_1\boldsymbol{z}_1^* \le 0$. Combining this with (10), we have that $\mathbf{L}_1\boldsymbol{x}_1 = 0$ and $-\tilde{U}(\boldsymbol{x}_1^*,\boldsymbol{x}_2^*) + \tilde{U}(\boldsymbol{x}_1,\boldsymbol{x}_2^*) = 0$. Since $\tilde{U}$ is strictly concave in its first argument $\boldsymbol{x}_1 = \boldsymbol{x}_1^*$. A similar argument establishes that $\boldsymbol{x}_2 = \boldsymbol{x}_2^*$. Using now the fact that M is weakly positively invariant, one can deduce that $\mathbf{L}_\ell\boldsymbol{z}_\ell \in -\partial_{\boldsymbol{x}_\ell}\tilde{U}(\boldsymbol{x}_1,\boldsymbol{x}_2)$, for $\ell \in \{1,2\}$, and thus $(\boldsymbol{x}_1,\boldsymbol{z}_1,\boldsymbol{x}_2,\boldsymbol{z}_2) \in \mathrm{Eq}(\Psi_{\text{Nash-undir}})$. $\square$

5 Distributed convergence to Nash equilibria for directed topologies

Interestingly, the saddle-point dynamics (8) fails to converge when transcribed to the directed network setting. This observation is a consequence of the following result, which studies the stability of the linearization of the dynamics (8), when the payoff functions have no contribution to the linear part.

Lemma 5.1 (Necessary condition for the convergence of (8) on digraphs): *Let Σ_ℓ be strongly connected and $f_\ell^i = 0$, $i \in \{1,\dots,n_\ell\}$, for $\ell \in \{1,2\}$. Then, the set of network agreement configurations $\mathcal{S}_{agree} = \{(\mathbf{1}_{n_1}\otimes x_1, \mathbf{1}_{n_1}\otimes z_1, \mathbf{1}_{n_2}\otimes x_2, \mathbf{1}_{n_2}\otimes z_2) \in (\mathbb{R}^{n_1d_1})^2 \times (\mathbb{R}^{n_2d_2})^2 \mid x_\ell, z_\ell \in \mathbb{R}^{d_\ell}, \ell \in \{1,2\}\}$, is stable under (8) iff, for any nonzero eigenvalue λ of the Laplacian L_ℓ, $\ell \in \{1,2\}$, one has $\sqrt{3}|\mathrm{Im}(\lambda)| \le \mathrm{Re}(\lambda)$.*

Proof. In this case, (8) is linear with matrix

$$\begin{pmatrix} \left(\begin{smallmatrix} -1 & -1 \\ 1 & 0 \end{smallmatrix}\right) \otimes \mathbf{L}_1 & 0 \\ 0 & \left(\begin{smallmatrix} -1 & -1 \\ 1 & 0 \end{smallmatrix}\right) \otimes \mathbf{L}_2 \end{pmatrix} \tag{11}$$

and has $\mathcal{S}_{\text{agree}}$ as equilibria. The eigenvalues of (11) are of the form $\lambda_\ell\left(\frac{-1}{2} \pm \frac{\sqrt{3}}{2}i\right)$, with λ_ℓ an eigenvalue of $\mathbf{L}_\ell$, for $\ell \in \{1,2\}$ (since the eigenvalues of a Kronecker product are the product of the eigenvalues of the corresponding matrices). Since $\mathbf{L}_\ell = \mathsf{L}_\ell \otimes \mathsf{I}_{d_\ell}$, each eigenvalue of $\mathbf{L}_\ell$ is an eigenvalue of L_ℓ. The result follows by noting that $\mathrm{Re}\big(\lambda_\ell(\frac{-1}{2} \pm \frac{\sqrt{3}}{2}i)\big) = \frac{1}{2}(\mp\sqrt{3}\mathrm{Im}(\lambda_\ell) - \mathrm{Re}(\lambda_\ell))$. $\square$

It is not difficult to construct examples of strictly concave-convex functions that have zero contribution to the linearization of (8) around the solution. Therefore, such systems cannot be convergent if they fail the necessary condition identified in Lemma 5.1.

The counterexample provided in our recent paper (Gharesifard and Cortés, 2012c) of strongly connected, weight-balanced digraphs that do not meet the stability criterium of Lemma 5.1 is therefore valid in this context too.

From here on, we assume that the payoff functions are differentiable. We elaborate on the reasons for this assumption in Remark 5.3 later. Motivated by the observation made in Lemma 5.1, we introduce a parameter $\alpha \in \mathbb{R}_{>0}$ in the dynamics of (8) as

$$\dot{\boldsymbol{x}}_1 + \alpha \mathbf{L}_1 \boldsymbol{x}_1 + \mathbf{L}_1 \boldsymbol{z}_1 = \nabla \tilde{U}(\boldsymbol{x}_1, \boldsymbol{x}_2), \tag{12a}$$

$$\dot{\boldsymbol{z}}_1 = \mathbf{L}_1 \boldsymbol{x}_1, \tag{12b}$$

$$\dot{\boldsymbol{x}}_2 + \alpha \mathbf{L}_2 \boldsymbol{x}_2 + \mathbf{L}_2 s \boldsymbol{z}_2 = -\nabla \tilde{U}(\boldsymbol{x}_1, \boldsymbol{x}_2), \tag{12c}$$

$$\dot{\boldsymbol{z}}_2 = \mathbf{L}_2 \boldsymbol{x}_2. \tag{12d}$$

We next show that a suitable choice of α makes the dynamics convergent to the Nash equilibrium.

Theorem 5.2 (Distributed convergence to Nash equilibria for directed networks): *Consider the zero-sum game* $\mathbf{G}_{\text{adv-net}} = (\Sigma_1, \Sigma_2, \Sigma_{\text{eng}}, U)$, *with* Σ_1 *and* Σ_2 *strongly connected and weight-balanced digraphs,* $\mathsf{X}_1 \subset \mathbb{R}^{d_1}$, $\mathsf{X}_2 \subset \mathbb{R}^{d_2}$ *compact and convex, and* $U : \mathsf{X}_1 \times \mathsf{X}_2 \to \mathbb{R}$ *strictly concave-convex and differentiable with globally Lipschitz gradient. Assume* $\mathbf{G}_{\text{adv-net}}$ *can be lifted to* $\tilde{\mathbf{G}}_{\text{adv-net}}$ *such that* $\tilde{U}$ *is differentiable and has a globally Lipschitz gradient. Define* $h : \mathbb{R}_{>0} \to \mathbb{R}$ *by*

$$h(r) = \frac{1}{2}\Lambda_*^{\min}\Big(\sqrt{\Big(\frac{r^4+3r^2+2}{r}\Big)^2 - 4} - \frac{r^4+3r^2+2}{r}\Big) + \frac{Kr^2}{(1+r^2)}, \tag{13}$$

where $\Lambda_*^{\min} = \min_{\ell=1,2}\{\Lambda_*(\mathsf{L}_\ell + \mathsf{L}_\ell^T)\}$, $\Lambda_*(\cdot)$ *denotes the smallest non-zero eigenvalue and* $K \in \mathbb{R}_{>0}$ *is the Lipschitz constant of the gradient of* $\tilde{U}$. *Then there exists* $\beta^* \in \mathbb{R}_{>0}$ *with* $h(\beta^*) = 0$ *such that for all* $0 < \beta < \beta^*$, *the projection onto the first and third components of the solutions of* (12) *with* $\alpha = \frac{\beta^2+2}{\beta}$ *asymptotically converge to agreement on the Nash equilibrium of* $\mathbf{G}_{\text{adv-net}}$.

Proof. Similarly to the proof of Theorem 4.1, we assume, without loss of generality, that agents in Σ_1 have access to $\boldsymbol{x}_2$ and agents in Σ_2 to $\boldsymbol{x}_1$. For convenience, we denote the dynamics described in (12) by $\Psi_{\text{Nash-dir}} : (\mathbb{R}^{d_1 n_1})^2 \times (\mathbb{R}^{d_2 n_2})^2 \to (\mathbb{R}^{d_1 n_1})^2 \times (\mathbb{R}^{d_2 n_2})^2$. Let $(\boldsymbol{x}_1^0, \boldsymbol{z}_1^0, \boldsymbol{x}_2^0, \boldsymbol{z}_2^0)$ be an arbitrary initial condition. Note that the set $W_{\boldsymbol{z}_1^0, \boldsymbol{z}_2^0}$ defined by (9) is invariant under the evolutions of (12). By an argument similar to the one in the proof of Theorem 4.1, there exists a unique solution to (7), which we denote by $\boldsymbol{x}_1^* = \mathbf{1}_{n_1} \otimes x_1^*$ and $\boldsymbol{x}_2^* = \mathbf{1}_{n_2} \otimes x_2^*$. By Proposition 3.4(i), there exists $(\boldsymbol{x}_1^*, \boldsymbol{z}_1^*, \boldsymbol{x}_2^*, \boldsymbol{z}_2^*) \in \text{Eq}(\Psi_{\text{Nash-dir}}) \cap W_{\boldsymbol{z}_1^0, \boldsymbol{z}_2^0}$. Consider the function $V : (\mathbb{R}^{d_1 n_1})^2 \times (\mathbb{R}^{d_2 n_2})^2 \to \mathbb{R}_{\geq 0}$,

$$\begin{aligned} V(\boldsymbol{x}_1, \boldsymbol{z}_1, \boldsymbol{x}_2, \boldsymbol{z}_2) &= \frac{1}{2}(\boldsymbol{x}_1 - \boldsymbol{x}_1^*)^T(\boldsymbol{x}_1 - \boldsymbol{x}_1^*) + \frac{1}{2}(\boldsymbol{x}_2 - \boldsymbol{x}_2^*)^T(\boldsymbol{x}_2 - \boldsymbol{x}_2^*) \\ &+ \frac{1}{2}(\boldsymbol{y}_{(\boldsymbol{x}_1, \boldsymbol{z}_1)} - \boldsymbol{y}_{(\boldsymbol{x}_1^*, \boldsymbol{z}_1^*)})^T(\boldsymbol{y}_{(\boldsymbol{x}_1, \boldsymbol{z}_1)} - \boldsymbol{y}_{(\boldsymbol{x}_1^*, \boldsymbol{z}_1^*)}), \\ &+ \frac{1}{2}(\boldsymbol{y}_{(\boldsymbol{x}_2, \boldsymbol{z}_2)} - \boldsymbol{y}_{(\boldsymbol{x}_2^*, \boldsymbol{z}_2^*)})^T(\boldsymbol{y}_{(\boldsymbol{x}_2, \boldsymbol{z}_2)} - \boldsymbol{y}_{(\boldsymbol{x}_2^*, \boldsymbol{z}_2^*)}), \end{aligned}$$

where $\boldsymbol{y}_{(\boldsymbol{x}_\ell, \boldsymbol{z}_\ell)} = \beta \boldsymbol{x}_\ell + \boldsymbol{z}_\ell$, $\ell \in \{1, 2\}$, and $\beta \in \mathbb{R}_{>0}$ satisfies $\beta^2 - \alpha\beta + 2 = 0$. This function is quadratic, hence smooth. Next, we consider $\xi = \mathcal{L}_{\Psi_{\text{Nash-dir}}} V(\boldsymbol{x}_1, \boldsymbol{z}_1, \boldsymbol{x}_2, \boldsymbol{z}_2)$ given by

$$\begin{aligned} \xi = (&-\alpha \mathbf{L}_1 \boldsymbol{x}_1 - \mathbf{L}_1 \boldsymbol{z}_1 + \nabla \tilde{U}(\boldsymbol{x}_1, \boldsymbol{x}_2), \mathbf{L}_1 \boldsymbol{x}_1, -\alpha \mathbf{L}_2 \boldsymbol{x}_2 \\ &- \mathbf{L}_2 \boldsymbol{z}_2 - \nabla \tilde{U}(\boldsymbol{x}_1, \boldsymbol{x}_2), \mathbf{L}_2 \boldsymbol{x}_2) \cdot \nabla V(\boldsymbol{x}_1, \boldsymbol{z}_1, \boldsymbol{x}_2, \boldsymbol{z}_2). \end{aligned}$$

After some manipulation, one can show that

$$\begin{aligned} \xi = &\sum_{\ell=1}^{2} \frac{1}{2}(\boldsymbol{x}_\ell - \boldsymbol{x}_\ell^*, \boldsymbol{y}_{(\boldsymbol{x}_\ell, \boldsymbol{z}_\ell)} - \boldsymbol{y}_{(\boldsymbol{x}_\ell^*, \boldsymbol{z}_\ell^*)})^T A_\ell(\boldsymbol{x}_l, \boldsymbol{y}_{(\boldsymbol{x}_\ell, \boldsymbol{z}_\ell)}) \\ &+ \sum_{\ell=1}^{2} \frac{1}{2}(\boldsymbol{x}_\ell^T, \boldsymbol{y}_{(\boldsymbol{x}_\ell, \boldsymbol{z}_\ell)}^T) A_\ell^T(\boldsymbol{x}_\ell - \boldsymbol{x}_\ell^*, \boldsymbol{y}_{(\boldsymbol{x}_\ell, \boldsymbol{z}_\ell)} - \boldsymbol{y}_{(\boldsymbol{x}_\ell^*, \boldsymbol{z}_\ell^*)}) \\ &+ \sum_{\ell=1}^{2} (-1)^{j-1}(\boldsymbol{x}_\ell - \boldsymbol{x}_\ell^*)^T \nabla_{\boldsymbol{x}_\ell} \tilde{U}(\boldsymbol{x}_1, \boldsymbol{x}_2) \\ &+ \sum_{\ell=1}^{2} (-1)^{j-1}\beta(\boldsymbol{y}_{(\boldsymbol{x}_\ell, \boldsymbol{z}_\ell)} - \boldsymbol{y}_{(\boldsymbol{x}_\ell^*, \boldsymbol{z}_\ell^*)})^T \nabla_{\boldsymbol{x}_\ell} \tilde{U}(\boldsymbol{x}_1, \boldsymbol{x}_2), \end{aligned}$$

where A_ℓ, $\ell \in \{1, 2\}$, is

$$A_\ell = \begin{pmatrix} -(\alpha - \beta)\mathbf{L}_\ell & -\mathbf{L}_\ell \\ (-\beta(\alpha - \beta) + 1)\mathbf{L}_\ell & -\beta \mathbf{L}_\ell \end{pmatrix}.$$

This equation can be written as

$$\begin{aligned} \xi = &\sum_{\ell=1}^{2} \frac{1}{2}(\boldsymbol{x}_\ell - \boldsymbol{x}_\ell^*, \boldsymbol{y}_{(\boldsymbol{x}_\ell, \boldsymbol{z}_\ell)} - \boldsymbol{y}_{(\boldsymbol{x}_\ell^*, \boldsymbol{z}_\ell^*)})^T \\ &\qquad Q_\ell(\boldsymbol{x}_\ell - \boldsymbol{x}_\ell^*, \boldsymbol{y}_{(\boldsymbol{x}_\ell, \boldsymbol{z}_\ell)} - \boldsymbol{y}_{(\boldsymbol{x}_\ell^*, \boldsymbol{z}_\ell^*)}) \\ &+ \sum_{\ell=1}^{2} (\boldsymbol{x}_\ell - \boldsymbol{x}_\ell^*, \boldsymbol{y}_{(\boldsymbol{x}_\ell, \boldsymbol{z}_\ell)} - \boldsymbol{y}_{(\boldsymbol{x}_\ell^*, \boldsymbol{z}_\ell^*)})^T A_\ell(\boldsymbol{x}_\ell^*, \boldsymbol{y}_{(\boldsymbol{x}_\ell^*, \boldsymbol{z}_\ell^*)}) \\ &+ \sum_{\ell=1}^{2} (-1)^{j-1}(\boldsymbol{x}_\ell - \boldsymbol{x}_\ell^*)^T \nabla_{\boldsymbol{x}_\ell} \tilde{U}(\boldsymbol{x}_1, \boldsymbol{x}_2) \\ &+ \sum_{\ell=1}^{2} (-1)^{j-1}\beta(\boldsymbol{y}_{(\boldsymbol{x}_\ell, \boldsymbol{z}_\ell)} - \boldsymbol{y}_{(\boldsymbol{x}_\ell^*, \boldsymbol{z}_\ell^*)})^T \nabla_{\boldsymbol{x}_\ell} \tilde{U}(\boldsymbol{x}_1, \boldsymbol{x}_2), \end{aligned}$$

where Q_ℓ, $\ell \in \{1,2\}$, is given by

$$Q_\ell = (\mathbf{L}_\ell + \mathbf{L}_\ell^T) \otimes \begin{pmatrix} -(\frac{\beta^2+2}{\beta} - \beta) & -1 \\ -1 & -\beta \end{pmatrix}. \qquad (14)$$

Note that, we have

$$\begin{aligned} A_1(\boldsymbol{x}_1^*, \boldsymbol{y}_{(\boldsymbol{x}_1^*,\boldsymbol{z}_1^*)}) &= -(\mathbf{L}_1 \boldsymbol{y}_{(\boldsymbol{x}_1^*,\boldsymbol{z}_1^*)}, \beta \mathbf{L}_1 \boldsymbol{y}_{(\boldsymbol{x}_1^*,\boldsymbol{z}_1^*)}) \\ &= -(\nabla_{\boldsymbol{x}_1} \tilde{U}(\boldsymbol{x}_1^*, \boldsymbol{x}_2), \beta \nabla_{\boldsymbol{x}_1} \tilde{U}(\boldsymbol{x}_1^*, \boldsymbol{x}_2)), \\ A_2(\boldsymbol{x}_2^*, \boldsymbol{y}_{(\boldsymbol{x}_2^*,\boldsymbol{z}_2^*)}) &= -(\mathbf{L}_2 \boldsymbol{y}_{(\boldsymbol{x}_2^*,\boldsymbol{z}_2^*)}, \beta \mathbf{L}_2 \boldsymbol{y}_{(\boldsymbol{x}_2^*,\boldsymbol{z}_2^*)}) \\ &= (\nabla_{\boldsymbol{x}_2} \tilde{U}(\boldsymbol{x}_1, \boldsymbol{x}_2^*), \beta \nabla_{\boldsymbol{x}_2} \tilde{U}(\boldsymbol{x}_1, \boldsymbol{x}_2^*)). \end{aligned}$$

Thus, after substituting for $\boldsymbol{y}_{(\boldsymbol{x}_\ell,\boldsymbol{z}_\ell)}$, we have

$$\begin{aligned} \xi = \sum_{\ell=1}^{2} & \frac{1}{2}(\boldsymbol{x}_\ell - \boldsymbol{x}_\ell^*, \boldsymbol{z}_\ell - \boldsymbol{z}_\ell^*)^T \tilde{Q}_\ell (\boldsymbol{x}_\ell - \boldsymbol{x}_\ell^*, \boldsymbol{z}_\ell - \boldsymbol{z}_\ell^*) \\ &+ (1+\beta^2)(\boldsymbol{x}_1 - \boldsymbol{x}_1^*)^T (\nabla_{\boldsymbol{x}_1} \tilde{U}(\boldsymbol{x}_1, \boldsymbol{x}_2) - \nabla_{\boldsymbol{x}_1} \tilde{U}(\boldsymbol{x}_1^*, \boldsymbol{x}_2)) \\ &- (1+\beta^2)(\boldsymbol{x}_2 - \boldsymbol{x}_2^*)^T (\nabla_{\boldsymbol{x}_2} \tilde{U}(\boldsymbol{x}_1, \boldsymbol{x}_2) - \nabla_{\boldsymbol{x}_2} \tilde{U}(\boldsymbol{x}_1, \boldsymbol{x}_2^*)) \\ &+ \beta(\boldsymbol{z}_1 - \boldsymbol{z}_1^*)^T (\nabla_{\boldsymbol{x}_1} \tilde{U}(\boldsymbol{x}_1, \boldsymbol{x}_2) - \nabla_{\boldsymbol{x}_1} \tilde{U}(\boldsymbol{x}_1^*, \boldsymbol{x}_2)) \\ &- \beta(\boldsymbol{z}_2 - \boldsymbol{z}_2^*)^T (\nabla_{\boldsymbol{x}_2} \tilde{U}(\boldsymbol{x}_1, \boldsymbol{x}_2) - \nabla_{\boldsymbol{x}_2} \tilde{U}(\boldsymbol{x}_1, \boldsymbol{x}_2^*)), \quad (15) \end{aligned}$$

where

$$\tilde{Q}_\ell = \begin{pmatrix} -\beta^3 - (\frac{\beta^2+2}{\beta}) - \beta & -(1+\beta^2) \\ -(1+\beta^2) & -\beta \end{pmatrix} \otimes (\mathbf{L}_\ell + \mathbf{L}_\ell^T),$$

for $\ell \in \{1,2\}$. Each eigenvalue of $\tilde{Q}_\ell$ is of the form

$$\tilde{\eta}_\ell = \lambda_\ell \frac{-(\beta^4 + 3\beta^2 + 2) \pm \sqrt{(\beta^4 + 3\beta^2 + 2)^2 - 4\beta^2}}{2\beta}, \qquad (16)$$

where λ_ℓ is an eigenvalue of $\mathsf{L}_\ell + \mathsf{L}_\ell^T$, $\ell \in \{1,2\}$. Using now Theorem A.1 twice, one for $(\boldsymbol{x}_1, \boldsymbol{x}_2)$, $(\boldsymbol{x}_1^*, \boldsymbol{x}_2)$, and another one for $(\boldsymbol{x}_1, \boldsymbol{x}_2)$, $(\boldsymbol{x}_1, \boldsymbol{x}_2^*)$, we have

$$\begin{aligned} (\boldsymbol{x}_1 - \boldsymbol{x}_1^*)^T (\nabla_{\boldsymbol{x}_1} \tilde{U}(\boldsymbol{x}_1, \boldsymbol{x}_2) - \nabla_{\boldsymbol{x}_1} \tilde{U}(\boldsymbol{x}_1^*, \boldsymbol{x}_2)) &\leq \\ -\frac{1}{K}\left(||\nabla_{\boldsymbol{x}_1} \tilde{U}(\boldsymbol{x}_1, \boldsymbol{x}_2) - \nabla_{\boldsymbol{x}_1} \tilde{U}(\boldsymbol{x}_1^*, \boldsymbol{x}_2)||^2\right)&, \\ -(\boldsymbol{x}_2 - \boldsymbol{x}_2^*)^T (\nabla_{\boldsymbol{x}_2} \tilde{U}(\boldsymbol{x}_1, \boldsymbol{x}_2) - \nabla_{\boldsymbol{x}_2} \tilde{U}(\boldsymbol{x}_1, \boldsymbol{x}_2^*)) &\leq \\ -\frac{1}{K}\left(||\nabla_{\boldsymbol{x}_2} \tilde{U}(\boldsymbol{x}_1, \boldsymbol{x}_2) - \nabla_{\boldsymbol{x}_2} \tilde{U}(\boldsymbol{x}_1, \boldsymbol{x}_2^*)||^2\right)&, \end{aligned}$$

where $K \in \mathbb{R}_{>0}$ is the Lipschitz constant of $\nabla \tilde{U}$. We thus conclude that

$$\begin{aligned} \xi \leq \sum_{\ell=1}^{2} & \frac{1}{2}(\boldsymbol{x}_\ell - \boldsymbol{x}_\ell^*, \boldsymbol{z}_\ell - \boldsymbol{z}_\ell^*)^T \tilde{Q}_\ell (\boldsymbol{x}_\ell - \boldsymbol{x}_\ell^*, \boldsymbol{z}_\ell - \boldsymbol{z}_\ell^*) \\ &- \frac{(1+\beta^2)}{K}(||\nabla_{\boldsymbol{x}_1} \tilde{U}(\boldsymbol{x}_1, \boldsymbol{x}_2) - \nabla_{\boldsymbol{x}_1} \tilde{U}(\boldsymbol{x}_1^*, \boldsymbol{x}_2)||^2 \\ &+ ||\nabla_{\boldsymbol{x}_2} \tilde{U}(\boldsymbol{x}_1, \boldsymbol{x}_2) - \nabla_{\boldsymbol{x}_2} \tilde{U}(\boldsymbol{x}_1, \boldsymbol{x}_2^*)||^2) \\ &+ \beta(\boldsymbol{z}_1 - \boldsymbol{z}_1^*)^T (\nabla_{\boldsymbol{x}_1} \tilde{U}(\boldsymbol{x}_1, \boldsymbol{x}_2) - \nabla_{\boldsymbol{x}_1} \tilde{U}(\boldsymbol{x}_1^*, \boldsymbol{x}_2)) \\ &- \beta(\boldsymbol{z}_2 - \boldsymbol{z}_2^*)^T (\nabla_{\boldsymbol{x}_2} \tilde{U}(\boldsymbol{x}_1, \boldsymbol{x}_2) - \nabla_{\boldsymbol{x}_2} \tilde{U}(\boldsymbol{x}_1, \boldsymbol{x}_2^*)). \end{aligned}$$

One can write this inequality as displayed in (17), where

$$\begin{aligned} X = (&\boldsymbol{x}_1 - \boldsymbol{x}_1^*, \boldsymbol{z}_1 - \boldsymbol{z}_1^*, \boldsymbol{x}_2 - \boldsymbol{x}_2^*, \boldsymbol{z}_2 - \boldsymbol{z}_2^*, \\ &\nabla_{\boldsymbol{x}_1} \tilde{U}(\boldsymbol{x}_1, \boldsymbol{x}_2) - \nabla_{\boldsymbol{x}_1} \tilde{U}(\boldsymbol{x}_1^*, \boldsymbol{x}_2), \\ &\nabla_{\boldsymbol{x}_2} \tilde{U}(\boldsymbol{x}_1, \boldsymbol{x}_2) - \nabla_{\boldsymbol{x}_2} \tilde{U}(\boldsymbol{x}_1, \boldsymbol{x}_2^*)). \end{aligned}$$

Since $(\boldsymbol{x}_1, \boldsymbol{z}_1, \boldsymbol{x}_2, \boldsymbol{z}_2) \in W_{\boldsymbol{z}_1^0,\boldsymbol{z}_2^0}$, we have $(\mathbf{1}_{n_\ell}^T \otimes \mathsf{I}_{d_\ell})(\boldsymbol{z}_\ell - \boldsymbol{z}_\ell^*) = \mathbf{0}_{d_\ell}$, $\ell \in \{1,2\}$, and hence it is enough to establish that $\mathbf{Q}$ is negative semidefinite on the subspace $\mathcal{W} = \{(v_1, v_2, v_3, v_4, v_5, v_6) \in (\mathbb{R}^{n_1 d_1})^2 \times (\mathbb{R}^{n_2 d_2})^2 \times \mathbb{R}^{n_1 d_1} \times \mathbb{R}^{n_2 d_2} \mid (\mathbf{1}_{n_1}^T \otimes \mathsf{I}_{d_1}) v_2 = \mathbf{0}_{d_1}, (\mathbf{1}_{n_2}^T \otimes \mathsf{I}_{d_2}) v_4 = \mathbf{0}_{n_2}\}$. Using the fact that $-\frac{1}{K}(1+\beta^2)\mathsf{I}_{n_\ell d_\ell}$ is invertible, for $\ell \in \{1,2\}$, we can express $\mathbf{Q}$ as

$$\mathbf{Q} = N \underbrace{\begin{pmatrix} \bar{Q}_1 & 0 & 0 & 0 \\ 0 & \bar{Q}_2 & 0 & 0 \\ 0 & 0 & -\frac{1}{K}(1+\beta^2)\mathsf{I}_{n_1 d_1} & 0 \\ 0 & 0 & 0 & -\frac{1}{K}(1+\beta^2)\mathsf{I}_{n_2 d_2} \end{pmatrix}}_{\mathbf{D}} N^T,$$

where $\bar{Q}_\ell = \tilde{Q}_\ell + \frac{K\beta^2}{(1+\beta^2)} \begin{pmatrix} 0 & 0 \\ 0 & \mathsf{I}_{n_\ell d_\ell} \end{pmatrix}$, $\ell \in \{1,2\}$, and

$$N = \begin{pmatrix} \mathsf{I}_{n_1 d_1} & 0 & 0 & 0 & 0 & 0 \\ 0 & \mathsf{I}_{n_1 d_1} & 0 & 0 & -\frac{\beta K}{1+\beta^2}\mathsf{I}_{n_1 d_1} & 0 \\ 0 & 0 & \mathsf{I}_{n_2 d_2} & 0 & 0 & 0 \\ 0 & 0 & 0 & \mathsf{I}_{n_2 d_2} & 0 & \frac{\beta K}{1+\beta^2}\mathsf{I}_{n_2 d_2} \\ 0 & 0 & 0 & 0 & \mathsf{I}_{n_1 d_1} & 0 \\ 0 & 0 & 0 & 0 & 0 & \mathsf{I}_{n_2 d_2} \end{pmatrix}.$$

Noting that $\mathcal{W}$ is invariant under N^T (i.e., $N^T \mathcal{W} = \mathcal{W}$), all we need to check is that the matrix $\mathbf{D}$ is negative semidefinite on $\mathcal{W}$. Clearly,

$$\begin{pmatrix} -\frac{1}{K}(1+\beta^2)\mathsf{I}_{n_1 d_1} & 0 \\ 0 & -\frac{1}{K}(1+\beta^2)\mathsf{I}_{n_2 d_2} \end{pmatrix}$$

$$\xi \leq \frac{1}{2} X^T \underbrace{\begin{pmatrix} Q_{1,11} & Q_{1,12} & 0 & 0 & 0 & 0 \\ Q_{1,21} & Q_{1,22} & 0 & 0 & \beta \mathsf{I}_{n_1 d_1} & 0 \\ 0 & 0 & Q_{2,11} & Q_{2,12} & 0 & 0 \\ 0 & 0 & Q_{2,21} & Q_{2,22} & 0 & -\beta \mathsf{I}_{n_2 d_2} \\ 0 & \beta \mathsf{I}_{n_1 d_1} & 0 & 0 & -\frac{(1+\beta^2)}{K} \mathsf{I}_{n_1 d_1} & 0 \\ 0 & 0 & 0 & -\beta \mathsf{I}_{n_2 d_2} & 0 & -\frac{(1+\beta^2)}{K} \mathsf{I}_{n_2 d_2} \end{pmatrix}}_{\mathbf{Q}} X, \tag{17}$$

is negative definite. On the other hand, for $\ell \in \{1,2\}$, on $(\mathbb{R}^{n_\ell d_\ell})^2$, 0 is an eigenvalue of $\tilde{Q}_\ell$ with multiplicity $2d_\ell$ and eigenspace generated by vectors of the form $(\mathbf{1}_{n_\ell} \otimes a, 0)$ and $(0, \mathbf{1}_{n_\ell} \otimes b)$, with $a, b \in \mathbb{R}^{d_\ell}$. However, on $\{(v_1, v_2) \in (\mathbb{R}^{n_\ell d_\ell})^2 \mid (\mathbf{1}_{n_\ell}^T \otimes \mathsf{I}_{d_\ell}) v_2 = \mathbf{0}_{d_\ell}\}$, 0 is an eigenvalue of $\tilde{Q}_\ell$ with multiplicity d_ℓ and eigenspace generated by vectors of the form $(\mathbf{1}_{n_\ell} \otimes a, 0)$. Moreover, on $\{(v_1, v_2) \in (\mathbb{R}^{n_\ell d_\ell})^2 \mid (\mathbf{1}_{n_\ell}^T \otimes \mathsf{I}_{d_\ell}) v_2 = \mathbf{0}_{d_\ell}\}$, the eigenvalues of $\frac{K\beta^2}{(1+\beta^2)} \begin{pmatrix} 0 & 0 \\ 0 & \mathsf{I}_{n_\ell d_\ell} \end{pmatrix}$ are $\frac{K\beta^2}{(1+\beta^2)}$ with multiplicity $n_\ell d_\ell - d_\ell$ and 0 with multiplicity $n_\ell d_\ell$. Therefore, using Weyl's theorem (Horn and Johnson, 1985, Theorem 4.3.7), we deduce that the nonzero eigenvalues of the sum $\bar{Q}_\ell$ are upper bounded by $\Lambda_*(\tilde{Q}_\ell) + \frac{K\beta^2}{(1+\beta^2)}$. Thus, the eigenvalues of $\bar{Q} = \begin{pmatrix} \bar{Q}_1 & 0 \\ 0 & \bar{Q}_2 \end{pmatrix}$ are upper bounded by $\min_{\ell=1,2}\{\Lambda_*(\bar{Q}_\ell)\} + \frac{K\beta^2}{(1+\beta^2)}$. From (16) and the definition of h in (13), we conclude that the nonzero eigenvalues of $\bar{Q}$ are upper bounded by $h(\beta)$. It remains to show that there exists $\beta^* \in \mathbb{R}_{>0}$ with $h(\beta^*) = 0$ such that for all $0 < \beta < \beta^*$ we have $h(\beta) < 0$. For $r > 0$ small enough, $h(r) < 0$, since $h(r) = -\frac{1}{2}\Lambda_*^{\min} r + O(r^2)$. Furthermore, $\lim_{r \to \infty} h(r) = K > 0$. Hence, using the Mean Value Theorem, we deduce the existence of β^*. Therefore we conclude that $\mathcal{L}_{\Psi_{\text{Nash-dir}}} V(\boldsymbol{x}_1, \boldsymbol{z}_1, \boldsymbol{x}_2, \boldsymbol{z}_2) \leq 0$. As a by-product, the trajectories of (8) are bounded. Consequently, all assumptions of the LaSalle Invariance Principle, cf. Theorem 2.4, are satisfied. This result then implies that any trajectory of (8) starting from an initial condition $(\boldsymbol{x}_1^0, \boldsymbol{z}_1^0, \boldsymbol{x}_2^0, \boldsymbol{z}_2^0)$ converges to the largest invariant set M in $S_{\Psi_{\text{Nash-dir}}, V} \cap \mathcal{W}_{\boldsymbol{z}_1^0, \boldsymbol{z}_2^0}$. Note that if $(\boldsymbol{x}_1, \boldsymbol{z}_1, \boldsymbol{x}_2, \boldsymbol{z}_2) \in S_{\Psi_{\text{Nash-dir}}, V} \cap W_{\boldsymbol{z}_1^0, \boldsymbol{z}_2^0}$, then $N^T X \in \ker(\bar{Q}) \times \{0\}$. From the discussion above, we know $\ker(\bar{Q})$ is generated by vectors of the form $(\mathbf{1}_{n_1} \otimes a_1, 0, 0, 0)$, $(0, 0, \mathbf{1}_{n_2} \otimes a_2, 0)$, $a_\ell \in \mathbb{R}^{d_\ell}$, $j \in \{1,2\}$, and hence $\boldsymbol{x}_\ell = \boldsymbol{x}_\ell^* + \mathbf{1}_{n_\ell} \otimes a_\ell$, $\boldsymbol{z}_\ell = \boldsymbol{z}_\ell^*$. Using the strict concavity-convexity, this then implies that $\boldsymbol{x}_\ell = \boldsymbol{x}_\ell^*$. Finally, for $(\boldsymbol{x}_1^*, \boldsymbol{z}_1, \boldsymbol{x}_2^*, \boldsymbol{z}_2) \in M$, using the positive invariance of M, one deduces that $(\boldsymbol{x}_1^*, \boldsymbol{z}_1, \boldsymbol{x}_2^*, \boldsymbol{z}_2) \in \text{Eq}(\Psi_{\text{Nash-dir}})$. □

Remark 5.3 (Assumptions on payoff function): Two observations are in order regarding the assumptions in Theorem 5.2 on the payoff function. First, the assumption that the payoff function has a globally Lipschitz gradient is not too restrictive given that, since the state spaces are compact, standard boundedness conditions on the gradient imply the globally Lipschitz condition. Second, we restrict our attention to differentiable payoff functions because locally Lipschitz functions with globally Lipschitz generalized gradients are in fact differentiable, see (Gharesifard and Cortés, 2012c, Proposition A.1). •

Remark 5.4 (Comparison with best-response dynamics): Using the gradient flow has the advantage of avoiding the cumbersome computation of the best-response map. This, however, does not come for free. There are concave-convex functions for which the (distributed) gradient flow dynamics, unlike the best-response dynamics, fails to converge to the saddle point, see (Feijer and Paganini, 2010) for an example. •

We finish this section with an example.

Example 5.5 (Distributed adversarial selection of signal and noise power via (12)**):** Recall the communication scenario described in Remark 3.1. Consider 5 channels, $\{\mathtt{ch}_1, \mathtt{ch}_2, \mathtt{ch}_3, \mathtt{ch}_4, \mathtt{ch}_5\}$, for which the network Σ_1 has decided that $\{\mathtt{ch}_1, \mathtt{ch}_3\}$ have signal power x_1 and $\{\mathtt{ch}_2, \mathtt{ch}_4\}$ have signal power x_2. Channel $\mathtt{ch}_5$ has its signal power determined to satisfy the budget constraint $P \in \mathbb{R}_{>0}$, i.e., $P - 2x_1 - 2x_2$. Similarly, the network Σ_2 has decided that $\mathtt{ch}_1$ has noise power y_1, $\{\mathtt{ch}_2, \mathtt{ch}_3, \mathtt{ch}_4\}$ have noise power y_2, and $\mathtt{ch}_5$ has noise power $C - y_1 - 3y_2$ to meet the budget constraint $C \in \mathbb{R}_{>0}$. We let $\boldsymbol{x} = (x^1, x^2, x^3, x^4, x^5)$ and $\boldsymbol{y} = (y^1, y^2, y^3, y^4, y^5)$, where $x^i = (x_1^i, x_2^i) \in [0, P]^2$ and $y^i = (y_1^i, y_2^i) \in [0, C]^2$, for each $i \in \{1, \dots, 5\}$.

The networks Σ_1 and Σ_2, which are weight-balanced and strongly connected, and the engagement topology Σ_{eng} are shown in Figure 2. Note that, according to this topology, each agent can observe the power employed by its adversary in its channel and, additionally, the agents in channel 2 can obtain information about the estimates of the opponent in channel 4 and vice versa. The payoff functions of the agents are given in Remark 3.1, where for simplicity we take $\sigma_i = \sigma_1$, for $i \in \{1, 3, 5\}$, and $\sigma_i = \sigma_2$, for $i \in \{2, 4\}$, with $\sigma_1, \sigma_2 \in \mathbb{R}_{>0}$.

This example fits into the approach described in Section 3.1 by considering the following extended payoff

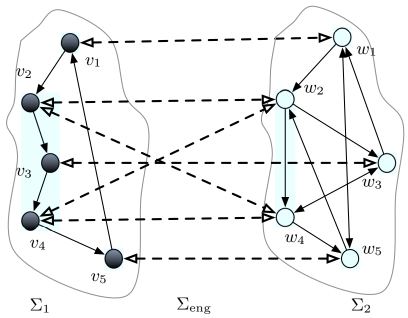


Fig. 2. Networks Σ_1, Σ_2 and Σ_{eng} for the case study of Example 5.5. Edges which correspond to Σ_{eng} are dashed. For $i \in \{1, \ldots, 5\}$, agents v_i and w_i are placed in channel $\mathtt{ch}_i$.

functions:

$$\begin{aligned}
\tilde{f}_1^1(x^1, \boldsymbol{y}) &= \log(1 + \frac{\beta x_1^1}{\sigma_1 + y_1^1}),\\
\tilde{f}_1^2(x^2, \boldsymbol{y}) &= \frac{1}{3}\log(1 + \frac{\beta x_2^2}{\sigma_2 + y_2^4}) + \frac{2}{3}\log(1 + \frac{\beta x_2^2}{\sigma_2 + y_2^2}),\\
\tilde{f}_1^3(x^3, \boldsymbol{y}) &= \log(1 + \frac{\beta x_1^3}{\sigma_1 + y_2^3}),\\
\tilde{f}_1^4(x^4, \boldsymbol{y}) &= \frac{1}{3}\log(1 + \frac{\beta x_2^4}{\sigma_2 + y_2^2}) + \frac{2}{3}\log(1 + \frac{\beta x_2^4}{\sigma_2 + y_2^4}),\\
\tilde{f}_1^5(x^5, \boldsymbol{y}) &= \log\Big(1 + \frac{\beta(P - 2x_1^5 - 2x_2^5)}{\sigma_1 + C - y_1^5 - 3y_2^5}\Big),\\
\tilde{f}_2^1(\boldsymbol{x}, y^1) &= \tilde{f}_1^1(x^1, \boldsymbol{y}), \quad \tilde{f}_2^3(\boldsymbol{x}, y^3) = \tilde{f}_1^3(x^3, \boldsymbol{y}),\\
\tilde{f}_2^2(\boldsymbol{x}, y^2) &= \frac{2}{3}\log(1 + \frac{\beta x_2^2}{\sigma_2 + y_2^2}) + \frac{1}{3}\log(1 + \frac{\beta x_2^4}{\sigma_2 + y_2^2}),\\
\tilde{f}_2^4(\boldsymbol{x}, y^4) &= \frac{1}{3}\log(1 + \frac{\beta x_2^2}{\sigma_2 + y_2^4}) + \frac{2}{3}\log(1 + \frac{\beta x_2^4}{\sigma_2 + y_2^4}),\\
\tilde{f}_2^5(\boldsymbol{x}, y^5) &= \tilde{f}_1^5(x^5, \boldsymbol{y}).
\end{aligned}$$

Note that these functions are strictly concave and thus the zero-sum game defined has a unique saddle point on the set $[0, P]^2 \times [0, C]^2$. These functions satisfy (5) and $\tilde{U}_1 = \tilde{U}_2$. Figure 3 shows the convergence of the dynamics (12) to the Nash equilibrium of the resulting 2-network zero-sum game. •

6 Conclusions and future work

We have considered a class of strategic scenarios in which two networks of agents are involved in a zero-sum game. The networks aim to either maximize or minimize a common objective function. Individual agents collaborate with neighbors in their respective network and have partial knowledge of the state of the agents in the other one. We have introduced two aggregate objective functions, one per network, identified a family of points with a special saddle property for this pair of functions, and established their correspondence between the Nash equilibria of the overall game. When the individual networks are undirected, we have proposed a distributed saddle-point dynamics that is implementable by each network via local interactions. We have shown that, for a class of strictly concave-convex and locally Lipschitz objective functions, the proposed dynamics is guaranteed to converge to the Nash equilibrium. We have also shown that this saddle-point dynamics fails to converge for directed networks, even when they are strongly connected and weight-balanced. Motivated by this fact, we have introduced a generalization that incorporates a design parameter. We have shown that this dynamics converges to the Nash equilibrium for strictly concave-convex and differentiable objective functions with globally Lipschitz gradients for appropriate parameter choices. An interesting venue of research is determining whether it is always possible to choose the extensions of the individual payoff functions in such a way that the lifted objective functions coincide. Future work will also include relaxing the assumptions of strict concavity-convexity and differentiability of the payoff functions, and the globally Lipschitz condition on their gradients, extending our results to dynamic interaction topologies and nonzero sum games, and exploring the application to various areas, including collaborative resource allocation in the presence of adversaries, strategic social networks, collective bargaining, and collaborative pursuit-evasion.

References

K. Arrow, L Hurwitz, and H. Uzawa. *A Gradient Method for Approximating Saddle Points and Constrained Maxima*. Rand Corporation, United States Army Air Forces, 1951.

K. Arrow, L Hurwitz, and H. Uzawa. *Studies in Linear and Non-Linear Programming*. Stanford University Press, Stanford, California, 1958.

E. N. Barron, R. Goebel, and R. R. Jensen. Best response dynamics for continuous games. *Proceeding of the American Mathematical Society*, 138(3):1069–1083, 2010.

T. Başar and G. J. Olsder. *Dynamic Noncooperative Game Theory*. SIAM, 2 edition, 1999. ISBN missing.

S. Boyd and L. Vandenberghe. *Convex Optimization*. Cambridge University Press, 2004. ISBN 0521833787.

F. Bullo, J. Cortés, and S. Martínez. *Distributed Control of Robotic Networks*. Applied Mathematics Series. Princeton University Press, 2009. ISBN 978-0-691-14195-4. Electronically available at http://coordinationbook.info.

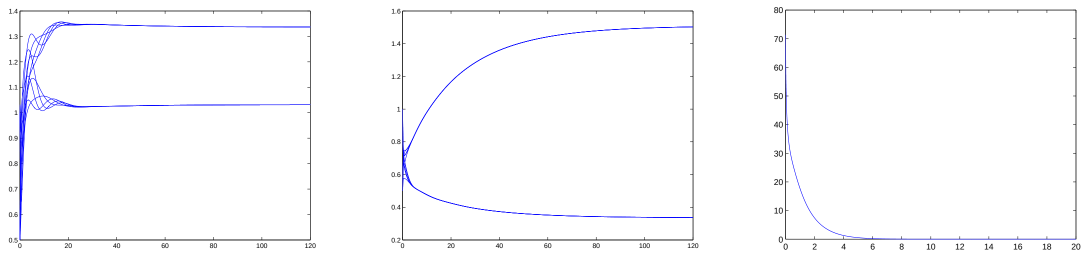
F. H. Clarke. *Optimization and Nonsmooth Analysis*. Canadian Mathematical Society Series of Monographs and Advanced Texts. Wiley, 1983. ISBN 047187504X.

J. Cortés. Discontinuous dynamical systems - a tutorial on solutions, nonsmooth analysis, and stability. *IEEE Control Systems Magazine*, 28(3):36–73, 2008.

D. Feijer and F. Paganini. Stability of primal-dual gradient dynamics and applications to network optimization. *Automatica*, 46:1974–1981, 2010.

P. Frihauf, M. Krstic, and T. Başar. Nash equilibrium seeking in noncooperative games. *IEEE Transactions on Automatic Control*, 2012. To appear.

B. Gharesifard and J. Cortés. Distributed convergence



(a) (b) (c)

Fig. 3. Execution of (12) over the networked strategic scenario described in Example 5.5, with $\beta = 8$, $\sigma_1 = 1$, $\sigma_2 = 4$, $P = 6$, and $C = 4$. (a) and (b) show the evolution of the agent's estimates of the state of networks Σ_1 and Σ_2, respectively, and (c) shows the value of the Lyapunov function. Here, $\alpha = 3$ in (12) and initially, $\boldsymbol{x}^0 = ((1, 0.5), (0.5, 1), (0.5, 0.5), (0.5, 1), (0.5, 1))^T$, $\boldsymbol{z}_1^0 = \mathbf{0}_{10}$, $\boldsymbol{y}^0 = ((1, 0.5), (0.5, 1), (0.5, 1), (0.5, 0.5), (1, 0.5))^T$ and $\boldsymbol{z}_2^0 = \mathbf{0}_{10}$. The equilibrium $(\boldsymbol{x}^*, \boldsymbol{z}_1^*, \boldsymbol{y}^*, \boldsymbol{z}_2^*)$ is $\boldsymbol{x}^* = (1.3371, 1.0315)^T \otimes \mathbf{1}_5$, $\boldsymbol{y}^* = (1.5027, 0.3366)^T \otimes \mathbf{1}_5$, $\boldsymbol{z}_1^* = (0.7508, 0.5084, 0.1447, 0.5084, 0.1447, -0.1271, -0.5201, -0.1271, -0.5201, -0.7626)^T$ and $\boldsymbol{z}_2^* = (0.1079, -0.0987, -0.0002, 0.2237, 0.0358, 0.2875, -0.0360, 0.0087, -0.1076, -0.4213)$.

to Nash equilibria by adversarial networks with undirected topologies. In *American Control Conference*, pages 5881–5886, Montréal, Canada, 2012a.

B. Gharesifard and J. Cortés. Distributed convergence to Nash equilibria by adversarial networks with directed topologies. In *IEEE Conf. on Decision and Control*, Maui, Hawaii, 2012b. To appear.

B. Gharesifard and J. Cortés. Continuous-time distributed convex optimization on directed graphs. *IEEE Transactions on Automatic Control*, 2012c. Conditionally accepted.

E. G. Golshtein and N. V. Tretyakov. *Modified Lagrangians and Monotone Maps in Optimization*. Wiley, New York, 1996.

J. Hofbauer and S. Sorin. Best response dynamics for continuous zero-sum games. *Discrete and Continuous Dynamical Systems Ser. B*, 6(1):215–224, 2006.

R. A. Horn and C. R. Johnson. *Matrix Analysis*. Cambridge University Press, 1985. ISBN 0521386322.

B. Johansson, M. Rabi, and M. Johansson. A randomized incremental subgradient method for distributed optimization in networked systems. *SIAM Journal on Control and Optimization*, 20(3):1157–1170, 2009.

J. S. Kim and S. Boyd. A minimax theorem with applications to machine learning, signal processing, and finance. *SIAM Journal on Optimization*, 19(3):1344–1367, 2008.

S. Li and T. Başar. Distributed algorithms for the computation of noncooperative equilibria. *Automatica*, 23 (4):523–533, 1987.

D. Maistroskii. Gradient methods for finding saddle points. *Matekon*, 13:3–22, 1977.

M. Mesbahi and M. Egerstedt. *Graph Theoretic Methods in Multiagent Networks*. Applied Mathematics Series. Princeton University Press, 2010.

A. Nedic and A. Ozdaglar. Distributed subgradient methods for multi-agent optimization. *IEEE Transactions on Automatic Control*, 54(1):48–61, 2009.

A. Nedic and A. Ozdgalar. Subgradient methods for saddle-point problems. *Journal of Optimization Theory & Applications*, 142(1):205–228, 2009.

R. Olfati-Saber, J. A. Fax, and R. M. Murray. Consensus and cooperation in networked multi-agent systems. *Proceedings of the IEEE*, 95(1):215–233, 2007.

W. Ren and R. W. Beard. *Distributed Consensus in Multi-vehicle Cooperative Control*. Communications and Control Engineering. Springer, 2008. ISBN 978-1-84800-014-8.

R. T. Rockafellar. *Convex Analysis*. Princeton Landmarks in Mathematics and Physics. Princeton University Press, Princeton, NJ, 1997. ISBN 0-691-01586-4. Reprint of 1970 edition.

M. S. Stankovic, K. H. Johansson, and D. M. Stipanovic. Distributed seeking of Nash equilibria with applications to mobile sensor networks. *IEEE Transactions on Automatic Control*, 2012. To appear.

P. Wan and M. D. Lemmon. Event-triggered distributed optimization in sensor networks. In *Symposium on Information Processing of Sensor Networks*, pages 49–60, San Francisco, CA, 2009.

J. Wang and N. Elia. Control approach to distributed optimization. In *Allerton Conf. on Communications, Control and Computing*, pages 557–561, Monticello, IL, October 2010.

J. Wang and N. Elia. A control perspective for centralized and distributed convex optimization. In *IEEE Conf. on Decision and Control*, pages 3800–3805, Orlando, Florida, 2011.

M. Zhu and S. Martínez. On distributed convex optimization under inequality and equality constraints. *IEEE Transactions on Automatic Control*, 57(1):151–164, 2012.

A Appendix

The following result can be understood as a generalization of the characterization of cocoercivity of concave functions (Golshtein and Tretyakov, 1996, Lemma 6.7).

Theorem A.1 (Concave-convex differentiable functions with globally Lipschitz gradients): *Let*

$f : \mathbb{R}^{d_1} \times \mathbb{R}^{d_2} \to \mathbb{R}$ be a concave-convex differentiable function with globally Lipschitz gradient (with Lipschitz constant $K \in \mathbb{R}_{>0}$). For $(x, y), (x', y') \in \mathbb{R}^{d_1} \times \mathbb{R}^{d_2}$,

$$\begin{aligned}
&(x - x')^T(\nabla_x f(x, y) - \nabla_x f(x', y')) \\
&\qquad + (y - y')^T(\nabla_y f(x', y') - \nabla_y f(x, y)) \\
&\leq -\frac{1}{2K}\Big(\|\nabla_x f(x, y') - \nabla_x f(x', y')\|^2 \\
&\quad + \|\nabla_y f(x', y) - \nabla_y f(x', y')\|^2 \\
&\quad + \|\nabla_x f(x', y) - \nabla_x f(x, y)\|^2 \\
&\quad + \|\nabla_y f(x, y') - \nabla_y f(x, y)\|^2\Big).
\end{aligned}$$

Proof. We start by noting that, for a concave function $j : \mathbb{R}^d \to \mathbb{R}$ with globally Lipschitz gradient, the following inequality holds, see (Golshtein and Tretyakov, 1996, Equation 6.64),

$$j(x) \leq j^* - \frac{1}{2M}||\nabla j(x)||^2, \tag{A.1}$$

where $j^* = \sup_{x \in \mathbb{R}^d} j(x)$ and M is the Lipschitz constant of ∇j. Given $(x', y') \in \mathbb{R}^{d_1} \times \mathbb{R}^{d_2}$, define the map $\tilde{f} : \mathbb{R}^{d_1} \times \mathbb{R}^{d_2} \to \mathbb{R}$ by

$$\begin{aligned}
\tilde{f}(x, y) = f(x, y) - f(x', y) - (x - x')^T \nabla_x f(x', y) \\
+ f(x, y) - f(x, y') + (y - y')^T \nabla_y f(x, y').
\end{aligned}$$

Since the gradient of f is Lipschitz, the function $\tilde{f}$ is differentiable almost everywhere. Thus, almost everywhere, we have

$$\begin{aligned}
\nabla_x \tilde{f}(x, y) &= \nabla_x f(x, y) - \nabla_x f(x', y) + \nabla_x f(x, y) \\
&\quad - \nabla_x f(x, y') - (y - y')^T \nabla_x \nabla_y f(x, y'), \\
\nabla_y \tilde{f}(x, y) &= \nabla_y f(x, y) - \nabla_y f(x', y) + \nabla_y f(x, y) \\
&\quad - \nabla_y f(x, y') - (x - x')^T \nabla_y \nabla_x f(x', y).
\end{aligned}$$

In particular, note that $\nabla_x \tilde{f}(x', y') = \nabla_y \tilde{f}(x', y') = 0$. Since $x \mapsto \tilde{f}(x, y')$ and $y \mapsto \tilde{f}(x', y)$ are concave and convex functions, respectively, we can use (A.1) to deduce

$$\tilde{f}(x, y') \leq -\frac{1}{2K}||\nabla_x f(x, y') - \nabla_x f(x', y')||^2, \tag{A.2a}$$

$$-\tilde{f}(x', y) \leq -\frac{1}{2K}||\nabla_y f(x', y) - \nabla_y f(x', y')||^2, \tag{A.2b}$$

where we have used the fact that $\sup_{x \in \mathbb{R}^{d_1}} \tilde{f}(x, y') = \inf_{y \in \mathbb{R}^{d_2}} \tilde{f}(x', y) = \tilde{f}(x', y') = 0$. Next, by definition of $\tilde{f}$,

$$\begin{aligned}
\tilde{f}(x, y') &= f(x, y') - f(x', y') - (x - x')^T \nabla_x f(x', y'), \\
\tilde{f}(x', y) &= f(x', y) - f(x', y') - (y - y')^T \nabla_y f(x', y').
\end{aligned}$$

Using (A.2), we deduce that

$$\begin{aligned}
&f(x, y') - f(x', y) \\
&- (x - x')^T \nabla_x f(x', y') + (y - y')^T \nabla_y f(x', y') \\
&\leq -\frac{1}{2K}\big(\|\nabla_x f(x, y') - \nabla_x f(x', y')\|^2 \\
&\qquad\qquad + \|\nabla_y f(x', y) - \nabla_y f(x', y')\|^2\big).
\end{aligned} \tag{A.3}$$

The claim now follows by adding together (A.3) and the inequality that results by interchanging (x, y) and (x', y') in (A.3). $\square$